EXAMPLE 5

High-level NPTII expression facilitates efficient recovery of transplastomic lines by selection for kanamycin resistance

The plastid genome of higher plants is a 120-kb to 160-kb double-stranded DNA which is present in 1,900 to 50,000 copies per leaf cell (Bendich, 1987). To obtain genetically stable transplastomic lines every one of the plastid genome copies (ptDNA) should be uniformly altered in a plant. Since integration of foreign DNA always occurs by homologous recombination, plastid transformation vectors contain segments of the plastid genome to target insertions at specific locations. Useful, non-selectable genes are cloned next to the selectable marker genes, which are then introduced into the plastid genome by linkage to the selectable marker gene (Maliga, 1993). Transforming DNA is introduced into plastids by the biolistic process (Svab et al., 1990; Svab and Maliga, 1993) or PEG treatment (Golds et al., 1993; O'Neil et al., 1993). Elimination of wild-type genome copies occurs during repeated cell divisions on a selective medium. The success of transformation depends on the success of selective amplification of the few initially transformed genome copies. Therefore the choice of the antibiotic used for the selective amplification of transformed genome copies and the mechanism by which the plant cells are protected from

antibiotic action is a critical parameter to be considered for successful generation of homoplasmic plants.

The most commonly used antibiotic for the selection of transplastomic lines is spectinomycin, an inhibitor of protein synthesis on plastid ribosomes. Initially, plastid transformation in tobacco was carried out by selection for resistance based on mutations in the plastid 16S rRNA (Svab et al., 1990). Selection was inefficient, yielding about one transplastomic clone per 50 bombarded samples, probably because the 16S rRNA based mutation in recessive. Recovery of transplastomic lines was enhanced ~100-fold by selection for a dominant marker, spectinomycin resistance based on inactivation by aminoglycoside 3" adenyltransferase encoded in a chimeric *aadA* gene (Svab and Maliga, 1993). In addition to tobacco, selection for spectinomycin resistance (*aadA*) could be applied to recover transplastomic lines in *Arabidopsis* and potato. The *aadA* gene in plants confers resistance to both spectinomycin and streptomycin. Selection for streptomycin resistance was used for plastid transformation in rice, a species resistant to spectinomycin, after bombardment with a chimeric *aadA* gene. See Example 8.

The need for an alternative marker gene for plastid manipulation has led to testing kanamycin resistance as a selective marker. A chimeric *neo* (*kan*) gene, encoding neomycin phosphotransferase, was suitable to recover transplastomic tobacco lines. However, recovery of transplastomic lines was relatively inefficient, yielding only one transplastomic line in ~25 bombarded leaf samples. Furthermore, for every plastid transformation event ~25 to 50 kanamycin resistant lines

were obtained in which integration of the plastid *neo* construct into the nuclear genome resulted in kanamycin resistance (Carrer *et al.*, 1993). We report here that the efficiency of recovering transplastomic clones is significantly improved when transforming tobacco chloroplasts with a new *neo* gene expressed from a promoter with the *atpB* and clpP translation control region. The number of nuclear transformation events is reduced using the cassettes of the present invention. These improvements make the new *neo* gene a practical tool for plastid genome manipulations.

DISCUSSION

The chimeric *neo* genes described in Examples 1-4 were introduced into plastids by selection for the linked spectinomycin resistance (*aadA*) gene as their suitability for directly selecting transplastomic lines was unknown. The transplastomic lines listed in Table 3 were then tested for resistance to kanamycin by their ability to proliferate on a medium containing 50 mg/L kanamycin. The RMOP meduim used for testing induces formation of green callus and shoot regeneration in the absence of kanamycin. The tissue culture procedures utilized for this example are described in references Carrer et al., 1993 and Carrer and Maliga, 1995.

On the selctive kanamycin medium only scanty, white callus forms from wild-type leaf section. Formation of green callus and shoots from leaf section of plants transformed with pHK plasmids in Table 3 indicates that accumulation of NPTII confers kanamycin resistance. We set out to test if transplastomic clones can be directly selected by kanamycin resistance after bombardment with

plasmids pHK30 and pHK32. The results are summarized in Table 5.

Bombardment of 25 tobacco leaves with plasmid pHK30 yielded 45 kanamycin resistant lines on a medium containing 50 mg/L kanamycin. Transplastomic *neo* lines are expected to be resistant to much higher levels, 500 mg/L of kanamycin (Carrer et al., 1993). In addition, in plasmid pHK30 the *neo* gene is physically linked to a spectinomycin resistance (*aadA*) gene. Spectinomycin resistance is manifested as kanamycin resistance: sensitive leaf sections form white callus and no shoots whereas resistant leaf sections form green callus and shoots on a selective medium (500 mg/L) RMOP medium. We assumed therefore, that all transplastomic lines should be resistant to both 500 mg/L of kanamycin and 500 mg/L spectinomycin (Carrer and Maliga, 1995). When applying this test we found that 22 of the 45 lines meet these criteria. Digestion of the plastid DNA with the EcoRI restriction enzyme and probing with the plastid targeting region should detect 3.1-kb fragment in the wild-type and a 4.2-kb and 1.2-kb fragment in transplastomic lines (Figure 15A). DNA gel blot analysis of seven of the kanamycin-spectinomycin resistant lines confirmed integration of both transgenes into the plastid genome (Figure 15B). Therefore, we assume that all 22 kanamycin-spectinomycin lines are transplastomic (Table 5).

Bombardment of 30 tobacco leaves with plasmid pHK32 yielded 28 kanamycin resistant lines on a medium containing 50 mg/L kanamycin. We have identified 11 double-resistant lines by testing these on a medium containing 500 mg/L of kanamycin and 500 mg/L spectinomycin. All six tested were transplastomic by DNA

gel blot analysis (Figure 15B), therefore we believe that all eleven are transplastomic (Table 5).

TABLE 5

SELECTION OF TRANSPLASTOMIC TOBACCO CLONES BY KANAMYCIN RESISTANCE

Vector	No. leaves	Kan. Res. 50 mg/L	Kan. Res. 500 mg/L	Kan. Res. 500 mg/L Spec. Res. 500 mg/L	Transplastomic
pTNH32	29	59	7		0
	50[a]	52			2
	25[a]	47	4		1
pHK30	25	45		22	22
pHK32	30	28		11	11

([a]Carrer et al., 1993)

DISCUSSION

Plastid transformation efficiency should be comparable, if we target the same region of the plastid genome for insertion, use similar size targeting sequences and the same method of DNA delivery. Therefore, lower transformation efficiencies obtained by selection for kanamycin resistance with the old chimeric *neo* genes was likely due to the lack of recovery of tranplastomic clones by selection . We have found that transformation with neo genes expressed from the

PrrnLatpB+DBwt and PrrnLclpP+DBwt promoters is as efficient as with the aadA gene. This is a significant technical advance, and will facilitate plastid transformation in crops, in which the regenerable tissues contain non-green plastids. Most important targets are the non-green plastids of cereal crops. Kanamycin selection is widely used to obtain transgenic lines after transformation with chimeric neo genes in dicots. However, kanamycin is an undesirable selective agent in monocots such as cereal tissue cultures. However, NPTII also inactivates paromomycin, which may be used to recover nuclear gene transformants at an extremely high efficiency in cereals. See for example, PCT application WO99/05296.

EXAMPLE 6

Bacterial *bar* gene expression in tobacco plastids confers resistance to the herbicide phosphinothricin

Bialaphos, a non-selective herbicide, is a tripeptide composed of two L-alanine residues and an analog of glutamic acid known as phosphinothricin (PPT). While PPT is an inhibitor of glutamine synthetase in both plants and bacteria, the intact tripeptide has little or no inhibitory effect in vitro. Bialaphos is toxic for bacteria and plants, as intracellular peptidases remove the alanine residues and release active PPT. Bialaphos is produced by *Streptomyces hygroscopicus*. The bacterium is protected from phosphinothricin toxicity by phosphinothricin acetyltransferase (PAT), the bar gene product. This enzyme acetylates phosphinothricin or demethylphosphinothricin (Thompson et al., 1987). PPT resistant crops have been obtained by expressing the *S.*

hygroscopicus bar gene in the plant nucleus. Herbicide resistant lines were obtained by direct selection for PPT resistance in culture after *Agrobacterium tumefaciens*-mediated DNA delivery in tobacco, potato, *Brassica napus* and *Brassica oleracea* (De Block et al., 1987, 1989). Biolistic DNA delivery of chimeric bar genes has been employed to obtain PPT resistant maize (Spencer et al., 1990), rice (Cao, et al, 1992) and Arabidopsis thaliana (Sawaskaki et al., 1994). Construction of transplastomic tobacco plants, in which PPT resistance is based on the expression of bar from S. hygroscopicus in plastids is described in the present example. The vectors utilized to express the bar gene contain an exemplary chimeric 5' regulatory region as set forth in the previous examples. The following material and methods facilitate the practice of this aspect of the present invention.

Construction of plastid bar gene

A NcoI/XbaI bar gene fragment was generated by PCR amplification using plasmid of pDM302 (Cao et al., 1992) with the following primers:

P1, 5'-AAACCATGGCACCACAAACAGAGAGCCCAGAACGACGCCC-3';

P2, 5'-AAAATCTAGATCATCAGATCTCGGTGACG-3'.

The ends of the PCR fragment were blunt ended by treatment with the Klenow Fragment of DNA polymerase I. The fragment was then ligated into the EcoRV site of pBluescript II KS+ (Stratagene, La Jolla, CA) to create plasmid pJEK3. Sequence analysis of pJEK3 plasmid DNA revealed that the XbaI site we intended to create through PCR amplification of pDM302 is absent. See Figure 19. The bar gene has the two translation

termination codons followed by vector sequences. The last 20 bp of pJEK3 are: CCCGTCACCGAGATCTGATGAtcgaattcctgcagcccgggggatccactagttct aga. The bar sequences are in capital (stop codons underlined), the vector sequences are in lower case (XbaI site underlined). Since there is an XbaI site present in the vector 40 bp from the intended XbaI site, it was not necessary to repair this error. The NcoI-XbaI fragment from plasmid pJEK3 was ligated into NcoI-XbaI digested pGS104 plasmid (Serino and Maliga, 1997) to generate plasmid pJEK6. Plasmid pGS104 carries a Prrn-TrbcL expression cassette in a pPRV111B plastid transformation vector. A map of the plastid targeting region of plasmid pJEK6 is shown in Figure 16A.

Plastid transformation and plant regeneration

Tobacco *(Nicotiana tabacum cv.* Petit Havana) plants were grown aseptically on agar-solidified medium containing MS salts (Murashige and Skoog, 1962) and sucrose (30g/l). Leaves were placed abaxial side up on RMOP media for bombardment. The RMOP medium consists of MS salts, N6-benzyladenine (1mg/l), 1-naphthaleneacetic acid (0.1 mg/l), thymine (1mg/l), inositol (100 mg/l), agar (6g/l), pH 5.8, and sucrose (30g/l). The DNA was introduced into chloroplasts on the surface of 1μm tungsten particles using the DuPont PDS1000He Biolistic gun (Maliga 1995). Spectinomycin resistant clones were selected on RMOP medium containing 500 μg/ml spectinomycin dihydrochloride. Resistant shoots were regenerated on the same selective medium and rooted on MS agar medium (Svab and Maliga, 1993). The independently transformed lines are designated by the

transforming plasmid (pJEK6) and a serial number, for example pJEK6-2, pJEK6-5. Plants regenerated from the same transformed line are distinguished by letters, for example pJEK6-2A, pJEK6-2B.

Southern Blot Analysis

Total cellular DNA was isolated from wild-type and transgenic spectinomycin resistant plants with CTAB (Saghai-Maroof et al., 1984). The DNA was digested with the Sma I and BglII restriction endonucleases, separated on a 0.7% agarose gel and blotted onto a Hybond-N nylon membrane (Amersham, Arlington Heights, IL) by a pressure blotter. The membrane was hybridized overnight with an ApaI/ BamHI fragment labeled with (α- ^{32}P)dCTP using a dCTP DNA Labeling Beads Kit (Pharmacia Inc, Piscataway, NJ). The membrane was washed 2 times with 0.1X SSPE, 0.2X SDS at 55°C for 30 minutes. Film was exposed to the membrane for 30 minutes at room temperature.

PAT Assay

The PAT assay was performed as described by Spencer et. al. (1990). Leaf tissue (100 mg) from wild type tobacco (wt), transgenic Nt-pDM307-10 tobacco (a line transformed with the nuclear bar gene in plasmid pDM307; Cao et al., 1992), and plastid bar gene transformants was homogenized in 1 volume of extraction buffer (10 mM Na_2HPO_4, 10 mM NaCl). The supernatant was collected after spinning in a microfuge for 10 minutes. Protein (25 mg) was added to 1 mg/ml PPT and ^{14}C-labeled Acetyl CoA. The reaction was incubated at 37°C for 30 minutes and the entire reaction was spotted onto a TLC plate. Ascending

chromatography was performed in a 3:2 mixture of 1-propanol and NH_4OH. Film was exposed to the TLC plate overnight at room temperature.

Herbicide Application

Wild type and transgenic plants were sprayed with 5 ml of a 2% solution of Liberty (AgrEvo, Wilmington, DE) with an aerosol sprayer.

RESULTS AND DISCUSSION

First the bacterial bar gene was converted into a plastid gene by cloning the bar coding region into a plastid expression cassette. This cassette consists of an engineered plastid rRNA operon promoter (Prrn) and TrbcL and the 3' UTR of the plastid rbcL gene for stabilization of the mRNA. The plastid bar gene was then cloned into the plastid transformation vector to yield plasmid pJEK6, and introduced into plastids on the surface of microscopic tungsten particles. The bar gene integrated into the plastid genome by two homologous recombination events via the plastid targeting sequences, as shown in Figure 16A. Selection for the linked aadA (spectinomycin resistance) gene on spectinomycin-containing medium eventually yielded cells which carried a uniformly transformed plastid genome population, which were then regenerated into plants.

Integration of bar and aadA was verified by DNA gel blot analysis. Total cellular DNA of wild-type and transplastomic plants was digested with the SmaI and

BglII restriction enzymes and probed with the 2.9-kb ApaI-BamHI plastid targeting fragment of *N. tabacum* (Figure 16B). The two fragments that were expected for the transgenic plants, 3.3 kb and 1.9 kb, were present in each of the transplastomic samples shown in Figure 16B. Absence of the 2.9 kb wild type fragment indicated, that by the time these plants have been regenerated, the wild-type plastid genome copies have been diluted out on the selective medium.

To determine if the plastid bar gene has been expressed, leaf extracts were assayed for phosphinothricin acetyltransferase (PAT) activity. Conversion of PPT into acetyl-PPT indicated PAT activity in each of the tested transplastomic lines. Data in Figure 17 are shown for the transplastomic lines Nt-pJEK6-2D, Nt-pJEK6-5A and Nt-pJEK6-13B. Interestingly, PAT activity was significantly (>>10-fold) higher when bar was expressed in the plastids, as compared to the bar gene expressed from the cauliflower mosaic virus 35S promoter in the nucleus of the Nt-pDM307-10 plant.

PAT expression confers resistance to PPT in tissue culture and in the greenhouse. When wild type leaf sections are grown in tissue culture, 10 mg/L PPT completely blocks callus proliferation. This same PPT concentration is suitable for the selection of nuclear transformants after bombardment with the nuclear bar construct in plasmid pDM307. Leaf sections of plants expressing bar in plastids show resistance in the presence of up to 100 mg/L PPT in the culture medium. We have tested PPT resistance in the greenhouse, spraying wild-type and transplastomic plants with Liberty, a commercial formulation of PPT, at the recommended field dose of 2%. As shown in Figure 18A, 13 days after the

treatment, the wild type plants were dead while the transgenic plants thrived. Since then the sprayed plants have flowered and set seed. Figure 18B shows maternal inheritance of PPT resistance. Lack of plastid pollen transmission results in a lack of herbicide resistance in progeny pollinated with transgenic pollen. The bacterial bar gene has a high G + C content (68.3%; Genbank Accession No. X17220), while plastid genes have a relatively high A + T content; for example the G + C content of the highly expressed psbA and rbcL genes is 42.7% and 43.7%, respectively (Genbank Accession No. Z00044). Differences in the G + C content are also reflected in the codon usage biases. Interestingly, data presented here indicate that expression of bar from *S. hygroscopicus* is sufficiently high to confer resistance to field levels of the non-selective herbicide PPT. Furthermore, the PAT enzyme levels obtained in the transplastomic lines are significantly higher than those observed in the nuclear transformant. Therefore, further improvement of the expression levels may be obtained by optimizing the codon usage for plastids as set forth in Example 7.

Advantages of incorporating bar in the plastid genome are containment of herbicide resistance due to the lack of pollen transmission in most crops. Furthermore, the lack of genetic segregation would simplify back-crossing for the introduction of herbicide resistance into additional breeding lines.

EXAMPLE 7

A Synthetic *bar* gene Improves Containment and Enhances Expression in Plastids

The bacterial *bar* gene was introduced into the tobacco plastid genome by transformation with plasmid pJEK6, as described above in Example 6. In plasmid pJEK6 *bar* is expressed in a cassette consisting of the Prrn(L)rbcL(S) promoter and TrbcL transcription terminator. This plasmid conferred PPT resistance to plants grown in the presence of PPT in the tissue culture medium, but direct selection for transformed lines was not possible. Although the PAT levels in homoplastomic leaves was high, the amount of PAT produced by the few pJEK6 *bar* copies during the early stage of plastid transformation was probably insufficient to protect the entire cell.

To improve *bar* expression in plastids a synthetic gene was created. The codon usage was modified to mimic that of the average tobacco photosynthetic plastid gene. Changing the codon usage lead to a lowered GC content characteristic of higher plant plastid genes. To assist with cloning, restriction enzyme recognition sequences were removed and added as necessary. Codon usage frequency in bacteria reflects relative tRNA abundance: frequent use of codons for rare tRNAs may significantly reduce translation efficiency. We hoped that differential codon usage in plastids and bacteria would reduce or prevent expression of the synthetic gene in bacteria, thereby reducing the danger of horizontal gene transfer to microorganisms. We also hoped that improved *bar* expression in our novel promoter cassettes will allow direct selection of plastid transformants on PPT-containing medium.

<u>Materials and Methods for Example 7</u>

Codon comparisons of photosynthetic (*rbcL*, *psaA*, *psaB*, *psaC*, *psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*) plastid genes were compiled using GCG (Genetics Computer Group, Madison, WI). DNA mutations were then introduced into the bacterial *bar* gene making its codon usage more similar to plastid genes, while removing several restriction enzyme sites that could interfere with cloning. See Figure 28. The synthetic *bar* gene (*s-bar*) was obtained by single-step assembly of the entire *s-bar* gene from 28 oligonucleotides (one 44 nt primer, one 30 nt primer and twenty-six 40 nt primers) using PCR (Stemmer et al., 1995). The top and bottom strands of the primers overlap with each other by 20 nucleotides. *Nco*I and *Nhe*I sites were added at the 5' end and a *Xba*I site was added at the 3' end through PCR amplification. To obtain the complete *s-bar* gene, a small aliquot of the assembly PCR product was amplified using primers 1A and 14B. Unchanged nucleotides are in upper case, altered nucleotides are in lower case in the primers listed below.

Primer 1A ccATGgctAGCCCAGAAaGAaGaCCGGCCGAtATtaGaCG
Primer 1B GCATaTCaGCtTCtGTaGCACGtCtaATaTCGGCCGGtCt
Primer 2A TGCtACaGAaGCtGAtATGCCaGCaGTtTGtACaATCGTt
Primer 2B CTTGTtTCtATaTAaTGGTTaACGATtGTaCAaACtGCtG
Primer 3A AACCAtTAtATaGAaACAAGtACaGTaAACTTtaGaACtG
Primer 3B tTCtTGaGGTTCtTGaGGtTCaGTtCtaAAGTTtACtGTa
Primer 4A AaCCtCAaGAACCtCAaGAaTGGACtGAtGAtCTaGTCCG
Primer 4B AaGGATAGCGCTCtCGtAGACGGACtAGaTCaTCaGTCCA
Primer 5A TCTaCGaGAGCGCTATCCtTGGCTtGTaGCaGAaGTtGAC
Primer 5B GCGATaCCaGCtACtTCaCCGTCaACtTCtGCtACaAGCC
Primer 6A GGtGAaGTaGCtGGtATCGCaTAtGCGGGCCCtTGGAAGG
Primer 6B CCAaTCaTAtGCaTTtCtTGCCTTCCAaGGGCCCGCaTAt
Primer 7A CAaGaAAtGCaTAtGAtTGGACaGCtGAaTCaACtGTtTA

Primer 7B GtTGaTGaCGtGGtGAaACGTAaACaGTtGAtTCaGCtGT
Primer 8A CGTtTCaCCaCGtCAtCAaCGtACaGGACTtGGtTCtACt
Primer 8B TTCAGtAGaTGtGTaTAtAGaGTaGAaCCaAGtCCtGTaC
Primer 9A CTaTAtACaCAtCTaCTGAAaTCttTGGAGGCACAaGGtT
Primer 9B aACAGCtACaACaCTCTTaAAaCCtTGTGCCTCCAaaGAt
Primer10A TtAAGAGtGTtGTaGCTGTtATaGGatTGCCtAAtGAtCC
Primer10B CtTCaTGCATGCGtACaCtTGGaTCaTTaGGCAatCCtAT
Primer11A aAGtGTaCGCATGCAtGAaGCtCTaGGATATGCtCCaaGa
Primer11B CCtGCaGCCCtCAaCATaCCtCttGGaGCATATCCtAGaG
Primer12A GGtATGtTGaGGGCtGCaGGtTTCAAaCAtGGaAACTGGC
Primer12B tTGCCAaAAACCtACaTCATGCCAGTTtCCaTGtTTGAAa
Primer13A ATGAtGTaGGTTTtTGGCAaCTtGAtTTCAGtCTaCCaGT
Primer13B GtAGaACtGGACGaGGaGGTACtGGtAGaCTGAAaTCaAG
Primer14A ACCtCCtCGTCCaGTtCTaCCaGTtACtGAGATCTGATGA
Primer14B tctagaTCATCAGATCTCaGTaACtG

The amplified *s-bar* coding region was then cloned into a pBSIIKS+ plasmid (Stratagene, La Jolla, CA) and sequenced (Figure 20A). The *s-bar* gene was cloned into cassettes with the chimeric PrrnLatpB+DBwt, PrrnLrbcL+DBwt and PrrnLT7g10+DB/Ec promoters. Table 6 sets forth the plasmids used in the practice of this example.

Table6. Plasmids with *bar* genes.

Plasmid	Promoter	*bar*	3'UTR	Vector
pKO5		synthetic (*s-bar*)		pBSIIKS+
pKO3	PrrnLatpB+DBwt	synthetic (*s-bar*)	TrbcL	pPRV111B
pKO8	PrrnLrbcL+DBwt	synthetic (*s-bar*)	TrbcL	pPRV111A
pKO17	PrrnLT7g10+DB/Ec	synthetic (*s-bar*)	TrbcL	pPRV111B
pKO12	PrrnLrbcL+DBwt	bacterial (*bar*)	TrbcL	pPRV111A

To provide a suitable cloning site at 3'-end of the bacterial *bar* gene, the EagI/BglII fragment of *s-bar* was replaced with the cognate fragment of the bacterial *bar* coding region. Such a bacterial *bar* gene is incorporated in plasmid pKO12 (Figure 21). In plasmid pKO12 the first 22 nucleotides of the bacterial *bar* coding region are replaced with nucleotides from the *s-bar*.

RESULTS

The engineered bacterial *bar* gene in pJEK6 is expressed both in *E. coli* and plants, as shown in the previous example. We were interested to test if modification of the codon affects expression of the *s-bar* gene in plastids and in *E. coli*. In *E. coli*, *s-bar* expression was determined by measuring PAT activity. Extracts were prepared from bacteria carrying plasmids pKO3 and pKO8 expressing *s-bar* from the PrrnLatpB+DBwt and PrrnLrbcL+DBwt promoters, respectively. The radioactive assay did not detect any activity, although extracts from bacteria transformed with plasmids pJEK6 and pKO12 carrying the bacterial *bar* genes gave strong signals (Figure 22A). In plasmid pKO12 the first 22 nucleotides of the bacterial *bar* coding region are replaced with nucleotides from the *s-bar*. Therefore, lack of expression from the *s-bar* in *E. coli* is not due to changes within the first 22 nucleotides.

The *s-bar* was also introduced into plastids by transformation with vector pKO3. Extracts were prepared from pKO3- and pJEK6-transformed tobacco plants, which carry the *s-bar* and *bar* genes, respectively. Extracts from both types of plants contained significant PAT

activity (Figure 22B). Therefore, the synthetic *bar* is expressed in plastids but not in *E. coli*.

Changing the *bar* gene codon usage abrogated expression of the gene in *E. coli*. This is likely due to the introduction of the rare AGA and AGG arginine codons in the *s-bar* coding region. The triplet frequency per thousand nucleotides for AGA and AGG is the lowest in *E. coli*, reflecting low abundance of the tRNA required for translation of these codons. The minor arginine $tRNA^{Arg(AGG/AGA)}$ has been shown to be a limiting factor in the bacterial expression of several mammalian genes. The coexpression of *ArgU* (*dnaY*) gene that encodes for $tRNA^{Arg(AGG/AGA)}$ resulted in high level production of the target protein (Makrides 1996). The bacterial *bar* gene has 14 arginine codons, none of which are the rare AGA/AGG codons. The *s-bar* gene has five of them, three of which are located within the first 25 codons. Therefore, the likely explanation for the lack of *s-bar* expression in *E. coli* is introduction of the rare AGA and AGG arginine codons in the *s-bar* coding region.

There are proteins, which are toxic to *E. coli* but their expression is desirable in plastid to which it is not toxic. Engineering of these proteins in *E. coli* poses a problem, since the commonly used PEP plastid promoters are active in *E. coli*, thus the gene will be transcribed and the mRNA translated. Incorporation of minor codons in the coding region will prevent translation of these proteins in *E. coli*. Particularly useful in this regard is conversion of arginine codons to AGA/AGG. If no arginine is present in the N-terminal region, an N-terminal fusion may be designed containing multiple AGA/AGG codons to prevent translation of the mRNA.

Plants under field conditions are associated with microbes living in the soil, on the leaves and inside the plants. Gene flow from plastids to these microorganisms has not been shown. However, it would be an added safety measure to incorporate codons in plastid genes, which are rare in the target microorganisms, but are efficiently translated in plastids. Incorporation of AGA/AGG codons into the selective marker genes and the genes of interest will prevent transfer of genes from plants to microbes, which lack the capacity to efficiently translate the AGA/AGG codons. In case of specific plant-microbe associations, based on differences in codon usage preferences genes could be designed which would be expressed in plastids but not in microbes.

Attempts to directly select transplastomic clones after bombardment with the *s-bar* constructs so far has failed. The *s-bar* coding region in Figure 20A contains frequent and rare codons in proportions characteristic of plastid genes. It is possible, that relatively rare codons in a specific context at a critical stage will prevent recovery of plastid transformation events. Examples for tissue-specific translation of mRNAs dependent on tRNA availability are known (Zhou et al., 1999). Therefore, we designed a second synthetic bar gene, *S2-bar*, containing only frequent codons (Figure 20B). Plastid transformation with the s2-bar will enable direct selection of plastid transformation events by PPT resistance.

EXAMPLE 8

FLUORESCENT ANTIBIOTIC RESISTANCE MARKER FOR FACILE IDENTIFICATION OF TRANSPLASTOMIC CLONES IN TOBACCO AND RICE

Plastid transformation in higher plants is accomplished through a gradual process, during which all the 300-10,000 plastid genome copies are uniformly altered. Antibiotic resistance genes incorporated in the plastid genome facilitate maintenance of transplastomes during this process. Given the high number of plastid genome copies in a cell, transformation unavoidably yields chimeric tissues, in which the transplastomic cells need to be identified and regenerated into plants. In chimeric tissue, antibiotic resistance is not cell autonomous: transplastomic and wild-type sectors both are green due to phenotypic masking by the transgenic cells. Novel genes encoding FLARE-S, a fluorescent antibiotic resistance enzyme conferring resistance to spectinomycin and streptomycin, which were obtained by translationally fusing aminoglycoside 3''-adenylyltransferase [AAD] with the *Aequorea victoria* green fluorescent protein (GFP) are provided in the present example. FLARE-S facilitates distinction of transplastomic and wild-type sectors in the chimeric tissue, thereby significantly reducing the time and effort required to obtain genetically stable transplastomic lines. The utility of FLARE-S to select for plastid transformation events was shown by tracking segregation of transplastomic and wild-type plastids in tobacco and rice plants after transformation with FLARE-S plastid vectors and selection for resistance to spectinomycin and streptomycin, respectively.

Plastid transformation vectors contain a selectable marker gene and passenger gene(s) flanked by homologous plastid targeting sequences (Zoubenko *et al.*, 1994), and are introduced into plastids by biolistic DNA delivery (Svab et *al.*, 1990; Svab and Maliga, 1993) or PEG

treatment (Golds *et al.*, 1993; Koop *et al.*, 1996; O'Neill *et al.*, 1993). The selectable marker genes may encode resistance to spectinomycin, streptomycin or kanamycin. Resistance to the drugs is conferred by the expression of chimeric *aadA* (Svab and Maliga, 1993) and *neo* (*kan*) (Carrer *et al.*, 1993) genes in plastids. These drugs inhibit chlorophyll accumulation and shoot formation on plant regeneration media. The transplastomic lines are identified by the ability to form green shoots on bleached wild-type leaf sections. Obtaining a genetically stable transplastomic line involves cultivation of the cells on a selective medium, during which the cells divide at least 16 to 17 times (Moll *et al.*, 1990). During this time wild type and transformed plastids and plastid genome copies gradually sort out. The extended period of genome and organellar sorting yields chimeric plants consisting of sectors of wild-type and transgenic cells (Maliga, 1993). In the chimeric tissue antibiotic resistance conferred by *aadA* or *neo* is not cell autonomous: transplastomic and wild-type sectors are both green due to phenotypic masking by the transgenic tissue. Chimerism necessitates a second cycle of plant regeneration on a selective medium. In the absence of a visual marker this is an inefficient process, involving antibiotic selection and identification of transplastomic plants by PCR or Southern probing. The feasibility of visual identification of transformed sectors greatly reduces the effort required to obtain homoplastomic clones.

The *Aequorea victoria* green fluorescent protein (GFP) is a visual marker, allowing direct imaging of the fluorescent gene product in living cells without the need for prolonged and lethal histochemical staining

procedures. Its chromophore forms autocatalytically in the presence of oxygen and fluoresces green when absorbing blue or UV light (Prasher *et al.*, 1992; Chalfie *et al.*, 1994; Heim *et al.*, 1994) (reviewed in ref. Prasher, 1995; Cubitt *et al.*, 1995; Misteli and Spector, 1997). The *gfp* gene was modified for expression in the plant nucleus by removing a cryptic intron, introducing mutations to enhance brightness and to improve GFP solubility (Pang *et al.*, 1996; Reichel *et al.*, 1996; Rouwendal *et al.*, 1997; Haseloff *et al.*, 1997; Davis and Vierstra, 1998). GFP was used to monitor protein targeting to nucleus, cytoplasm and plastids from nuclear genes (Sheen *et al.*, 1995; Chiu *et al.*, 1996; Kšhler *et al.*, 1997), and to follow virus movement in plants (Baulcombe *et al.*, 1995; Epel *et al.*, 1996). GFP has also been used to detect transient gene expression in plastids (Hibberd *et al.*, 1998).

The expression of GFP by directly incorporating the *gfp* gene in the plastid genome is described herein. Incorporation of a visual marker, the GFP protein, in the plastid transformation vectors of the present invention facilitates distinction of spontaneous antibiotic resistant mutants and plastid transformants (Svab *et al.*, 1990). Furthermore, transplastomic sectors in the chimeric tissue can be visually identified, significantly reducing the time and effort required for obtaining genetically stable transplastomic lines. The utility of the GFP marker described here is further enhanced by its fusion with the enzyme aminoglycoside 3''-adenylyltransferase [AAD] conferring spectinomycin and streptomycin resistance to plants. Using a marker gene encoding a bifunctional protein, FLARE-S (<u>fl</u>uorescent <u>a</u>ntibiotic <u>r</u>esistance <u>e</u>nzyme, <u>s</u>pectinomycin

and streptomycin), prevents physical separation of the two genes and simplifies engineering. Furthermore, fluorescent antibiotic resistance genes enables extension of plastid transformation to cereal crops, in which plastid transformation is not associated with a readily identifiable tissue culture phenotype.

The following protocols are provided to facilitate the practice of the present example.

Construction of tobacco plastid vectors. The *aadA16gfp* gene encodes FLARE16-S fusion protein, and can be excised as an NheI-XbaI fragment from plasmid pMSK51, a pBSKSII+ derivative (Genbank Accesssion No. Not yet assigned . The fusion protein was obtained by cloning *gfp* (from plasmid pCD3-326F) downstream of *aadA* (in plasmid pMSK38), digesting the resulting plasmid with BstXI (at the 3' end of the *aadA* coding region) and NcoI (including the *gfp* translation initiation codon) and linking the two coding regions by a BstXI-NcoI compatible adapter. The adapter was obtained by annealing oligonucleotides 5'-GTGGGCAAAGAACTTGTTGAA GGAAAATTGGAGCTAGTAGAAGGTCTTAAAGTCGC-3' and 5'-CATGGCGACTTTAAGACCTTCTACTAGCTCCAATTTTCCTTCAACAAGTTCTTTGC CCACTACC-3'. The adapter connects AAD and GFP with a peptide of 16 amino acid residues (ELVEGKLELVEGLKVA).

The engineered *aadA* gene (Chinault *et al.*, 1986) in plasmid pMSK38 (pBSIIKS+ derivative) has NcoI and NheI sites at the 5' end and BstXI and XbaI sites at the 3' end of the gene. The NcoI site includes the translation initiation codon; the NheI and BstXI sites are in the coding region close to the 5' and 3' ends, respectively; the XbaI site is downstream of stop codon. The mutations were introduced by PCR using

oligonucleotides 5'-GGCCATGGGGGCTAGCGAAGCGGTGATCGCCGAAGTATCG-3' and 5'-CGAATTCTAGACATTATTTGCCCACTACCTTGGTGATCTC-3'.

The *gfp* gene in plasmid CD3-326F is the derivative of plasmid psmGFP, encoding the soluble modified version of GFP (accession number U70495) obtained under order number CD3-326 from the Arabidopsis Biological Resource Center, Columbus, OH (Davis and Vierstra, 1998). The *gfp* gene in plasmid CD3-326F is expressed in the PpsbA /TpsbA expression cassette. The *gfp* gene in plasmid CD3-326F was obtained through the following steps. The BamHI-SacI fragment from CD3-326 was cloned into pBSKS+ vector to yield plasmid CD3-326A. The SacI site downstream of the coding region was converted into an XbaI site by blunting and linker ligation (5'-GCTCTAGAGC; plasmid CD3-326B). An NcoI site was created to include the translation initiation codon and at the same time the internal NcoI site was removed by PCR amplification of the coding region N-terminus with primers 5'-CCGGATCCAAGGAGATATAACACCATGGCTAGTAAAGGAGAAGAACTTTTC-3' and 5'-GTGTTGGCCAAGGAACAGGTAGTTTTCC-3'. The PCR-amplified fragment was digested with BamHI and MscI restriction enzymes, and the resulting fragment was used to replace the BamHI-MscI fragment in plasmid CD3-326B to yield plasmid CD3-326C. The *gfp* coding region was excised from plasmid CD3-326C as an NcoI-XbaI fragment and cloned into a *psbA* cassette to yield plasmid CD3-326D. PpsbA and TpsbA are the *psbA* gene promoter and 3'- untranslated region derived from plasmids pJS25 (Staub and Maliga, 1993). TpsbA has been truncated by inserting a HindIII linker downstream of the modified BspHI site (Peter Hajdukiewcz, unpublished). The

PpsbA::gfp::TpsbA gene was excised as an EcoRI-HindIII fragment and cloned into EcoRI and HindIII digested pPRV111A, to yield plasmid CD3-326F.

The *aadA16gfp* coding region from plasmid pMSK51 was introduced into two expression cassettes. In plasmid pMSK53 the *aadA16gfp* coding region is expressed in the PrrnLrbcL+DBwt/TpsbA cassette, and encodes the FLARE16-S2 protein (fluorescent antibiotic resistance enzyme, spectinomcyin). PrrnLrbcL+DBwt is described in the previous examples and derives from plasmid pHK14. The construct contains a chimeric promoter composed of the *rrn* operon promoter, the *rbcL* gene leader and downstream box sequence. TpsbA is the *psbA* gene 3' untranslated region, and functions to stabilize the chimeric mRNA. In plasmid pMSK54 the *aadA16gfp* coding region is expressed in the PrrnLatpB+DBwt/TpsbA cassette, and encodes the FLARE16-S1 protein. PrrnLatpB+DBwt derives from plasmid pHK10, and is a chimeric promoter composed of the *rrn* operon promoter, the *atpB* leader and downstream box sequence. See Examples 1-4.

The chimeric *aadA16gfp* genes were introduced into the tobacco plastid transformation vector pPRV111B (Zoubenko *et al.*, 1994). The *aadA* gene was excised from plasmid pPRV111B with EcoRI and SpeI restriction enzymes, and replaced with the EcoRI-SpeI fragment from plasmids pMSK53 and pMSK54 to generate plasmids pMSK57 (*aadA16gfp-S2*) and pMSK56 (*aadA16gfp-S1*).

Construction of rice plastid vectors. Plasmid pMSK49 is a rice-specific plastid transformation vector which carries the *aadA11gfp-S3* gene as the selective marker in the *trnV/rps12/7* intergenic region (GenBank Accession Number: Not yet assigned). Plasmid pMSK49

carries the rice SmaI-SnaBI plastid fragment (restriction sites at nucleotides 122488 and 125 878 in the genome Hiratsuka *et al.*, 1989) cloned into a pBSKSII+ (Stratagene) vector after blunting the SacI and KpnI restriction sites. The XbaI site present in the rice plastid DNA fragment (position at nucleotide 125032 in the genome (Hiratsuka *et al.*, 1989) was removed by filling in and religation. Prior to cloning the selective marker the progenitor plasmid was digested with the BglII restriction enzyme giving rise to a deletion of 119 nucleotides between two proximal BglII sites (positions at 124367 and 124491). The *aadA11gfp-S3* gene was then cloned in the blunted BglII sites.

The *aadA* gene in plasmid pMSK49 was obtained by modifying the *aadA* gene in plasmid pMSK38 (above) to obtain plasmid pMSK39. The modification involved translationally fusing the *aadA* gene product at its N-terminus with an epitope of the human c-Myc protein (amino acids 410-419; EQKLISEEDL Kolodziej and Young, 1991). The genetic engineering was performed by ligating an adapter obtained by annealing complementary oligonucleotides with appropriate overhangs into NcoI-NheI digested pMSK38 plasmid. The oligonucleotides were: 5'-CATGGGGGCTAGCGAACAAAAACTCATTTCTGAAGAAGACTTGc-3' and 5'-CTAGGCAAGTCTTCTTCAGAAATGAGTTTTTGTTCGCTAGCCCC-3'.

The *aadA11gfp* gene encoding FLARE11-S was obtained by linking AAD and GFP with the 11-mer peptide ELAVEGKLEVA. To clone *aadA* and *gfp* in the same polycloning site, *gfp* (EcoRI-HindIII fragment; from plasmid CD3-326F) was cloned downstream of *aadA* in plasmid pMSK39 to obtain plasmid pMSK41. The two genes were excised together as an NheI-HindIII fragment, and cloned into plasmid pMSK45 to replace a kanamycin-

resistance gene yielding plasmid pMSK48. Plasmid pMSK45 is a derivative of plasmid pMSK35 which carries the PrrnLT7g10+DB/Ec promoter. The promoter consists of the plastid rRNA operon promoter and the leader sequence of the T7 phage gene 10 leader. In plasmid pMSK48, *aadA* is expressed from the PrrnLT7g10+DB/Ec promoter . The *aadA* and *gfp* genes were then translationally fused with an BstXI-NcoI adapter that links the AAD and GFP with an 11-mer peptide. The adapter was obtained by annealing oligonucleotides 5'-GTGGGCAAAGAACTTGCAGTTGAAGGAAAATTGGAGGTCGC-3' and 5'-CATGGCGACCTCCAATTTTCCTTCAACTGCAAGTTCTTTGCCCACTACC-3', which was ligated into BstXI/NcoI digested pMSK48 plasmid DNA to yield plasmid pMSK49. Plasmid pMSK49 has the rice plastid targeting sequences present in plasmid pMSK35.

Tobacco plastid transformation. Tobacco leaves from 4 to 6 weeks old plants were bombarded with DNA-coated tungsten particles using the Dupont PDS1000He Biolistic gun (1100 psi). Transplastomic clones were identified as green shoots regenerating on bleached leaf sections on RMOP medium containing 500mg/L spectinomycin dihydrochloride (Svab abd Maliga, 1993). The spectinomycin resistant shoots were illuminated with UV light (Model B 100AP, UV Products, Upland, California, USA). Shoots emitting green light were transferred to spectinomycin free MS medium (Murashige and Skoog, 1962) (3% sucrose) on which fluorescent (transplastomic) and non- fluorescent (wild-type) sectors formed. Fluorescent sectors were excised, and transferred to selective (500 mg/L spectinomycin) shoot regeneration (RMOP) medium. Regenerated shoots were tested for uniform transformation by Southern analysis.

Rice plastid transformation. Callus formation from mature *Oryza sativa* cv. Taipei 309 seeds was induced on a modified CIM medium (Tompson *et al.*, 1986), containing MS salts and vitamins (2 mg/L glycine, 0.5 mg/L nicotinic acid, 0.5 mg/L pyridoxine and 0.1 mg/L thiamine), 2 mg/L 2,4D, 1 mg/L kinetin and 300 mg/L casein enzymatic hydrolysate Type III (Sigma C-1026) and sucrose (30g/L). Embryogenic suspensions from the proliferating embryogenic calli were obtained on the AA medium (Muller and Grafe, 1978). For plastid transformation by the biolistic process rice embryogenic cells were plated on a filter paper on non-selective modified CIM medium (Tompson *et al.*, 1986). The bombarded cells were incubated for 48 hours, transferred to selective liquid AA medium (Muller and Grafe, 1978) (one to two weeks), and then to solid modified RRM regeneration medium (Zhang and Wu, 1988) containing MS salts and vitamins, 100 mg/L myo-inositol, 4 mg/L BAP, 0.5 mg/L IAA, 0.5 mg/L NAA, 30 g/L sucrose and 40 g/L maltose and 100 mg/L streptomycin sulfate on which green shoots appeared in two to three weeks. The shoots were rooted on a selective MS salt medium (Murashige and Skoog, 1962) containing 30 g/L sucrose and 100 mg/L streptomycin sulfate. Leaf samples for PCR analysis and confocal microscopy were taken from plants on selective medium.

PCR amplification of border fragments. Total cellular DNA was extracted according to Mettler (Mettler, 1987). The PCR analysis was carried out with a 9:1 mixture of AmpliTaq (Stratagene) and Vent (New England Biolabs) DNA polymerases in the Vent buffer following the manufacturer's recommendations. The left

border fragment was amplified with primers O3 (5'-ATGGATGAACTATACAAATAAG-3'and O4 (5'-GCTCCTATAGTGTGACG-3'). The right border fragment was amplified with primers O5 (5'-ACTACCTCTGATAGTTGAGTCG-3') and O6 (5'-AGAGGTTAATCGTACTCTGG-3'). The aadA part of FLARE-S genes was amplified with primers O1 (5'-GGCTCCGCAGTGGATGGCGGCCTG-3') and O2 (5'-GGGCTGATACTGGGCCGGCAGG-3'). Primer positions are shown in Fig. 5A. Note that the same primers can be used in transplastomic tobacco and rice plants expressing FLARE-S.

Detection of FLARE-S by fluorescence. FLARE-S expressing sectors in the leaves were visualized by an Olympus SZX stereo microscope equipped for GFP detection with a CCD camera system. Subcellular localization of GFP was verified by laser-scanning confocal microscopy (Sarastro 2000 Confocal Image System, Molecular Dynamics, Sunnyvale, CA). This system includes an argon mixed gas laser with lines at 488 and 568 nm and detector channels. The channels are adjusted for fluorescein and rhodamine images. GFP fluorescence was detected in the FITC channel (488-514 nm). Chlorophyll fluorescence was detected in the TRITC channel (560-580 nm). The images produced by GFP and chlorophyll fluorescence were viewed on a computer screen attached to the microscope and processed using the Adobe PhotoShop software.

Immunoblot analysis. Leaves (0.5 g) collected from plants in sterile culture were frozen in liquid nitrogen and ground to a fine powder in a mortar with a pestle. For protein extraction the powder was transferred to a

centrifuge tube containing 1 ml buffer [50 mM Hepes/KOH (pH 7.5), 1 mM EDTA, 10 mM potassium acetate, 5 mM magnesium acetate, 1 mM dithiothreitol and 2 mM PMSF] and mixed by flicking. The insoluble material was removed by centrifugation at 4°C for 5 min at 11,600 *g*. Protein concentration in the supernatant was determined using the Biorad protein assay reagent kit. Proteins (20 μl per lane) were separated in 12% SDS-PAGE (Laemmli, 1970). Proteins separated by SDS-PAGE were transferred to a Protran nitrocellulose membrane (Schleicher and Schuell) using a semi-dry electroblotting apparatus (Bio-Rad). The membrane was incubated with Living Colors Peptide Antibody (Clontech) diluted 1 to 200. FLARE-S was visualized using ECL chemilluminescence immunoblot detection on X-ray film. FLARE-S on the blots was quantified by comparison with a dilution series of commercially available purified wild-type GFP (Clontech).

RESULTS AND DISCUSSION

Tobacco plastid vectors with FLARE-S as the selectable marker.

Two FLARE-S fusion proteins were tested in *E. coli*. In one, the AAD and GFP were linked by an 11-mer (ELAVEGKLEVA), in the second by a 16-mer (ELVEGKLELVEGLKVA) linker. For transformation in tobacco, the *aadA16gfp* coding region (16-mer linker) was expressed in two cassettes known to mediate high levels of protein accumulation in plastids. Both utilize the strongest known plastid promoter driving the expression of the ribosomal RNA operon (Prrn), and the 3'-UTR of the highly expressed *psbA* gene (TpsbA) for the stabilization of the chimeric mRNAs. The PrrnLatpB+wtDB (plasmid pMSK56) and PrrnLrbcL+DBwt (plasmid pMSK57)

promoters utilize the *atpB* or *rbcL* gene leader sequences and the coding region N-termini with the downstream box (DB) sequence, respectively. Due to inclusion of the DB sequence in the chimeric genes, the proteins encoded by the two genes are slightly different, having 14 amino acids of the ATP-ase β subunit (*atpB* gene products) or ribulose 1,5-bisphosphate carboxylase/oxygenase (*rbcL* gene product) translationally fused with FLARE16-S (FLARE16-S1 and FLARE16-S2, respectively). To obtain a plastid transformation vector with the fluorescent spectinomycin resistance genes, the chimeric genes were cloned into the *trnV/rps12/7* plastid intergenic region in plastid vector pPRV111B. Plasmids pMSK56 and pMSK57 (Fig. 23) express FLARE16-S1 and FLARE16-S2, respectively, as markers.

Identification of transplastomic tobacco clones by fluorescence. Transformation was carried out by biolistic delivery of pMSK56 and pMSK57 plasmid DNA into chloroplast. The bombarded leaves were transferred onto selective (500 mg/L spectinomycin) shoot regeneration medium. Wild-type leaves on this medium bleach and form white callus. Cells with transformed plastids regenerate green shoots. The leaves on the selective medium were regularly inspected with a hand-held long-wave UV lamp for FLARE-S fluorescence.

No fluorescence could be detected in young shoots (3 to 5 mm in size) developing on pMSK56-bombarded leaves. However, formation of bright sectors in the leaves was observed, when these small shoots were transferred onto non-selective plant maintenance medium. In contrast, cultures bombarded with plasmid pMSK57 yielded small fluorescent shoots at an early stage.

These fluorescent shoots, and some of the non-fluorescent ones, developed into plants with bright sectors on non-selective plant maintenance medium. Therefore, FLARE16-S2 is useful for early detection of plastid transformation events. FLARE16-S2 fluorescence in young shoots on a selective medium should be due to relatively high levels of FLARE16-S2. Higher levels of FLARE16-S2 are also indicated by the brighter sectors in variegated leaves expressing FLARE16-S2 as compared to FLARE16-S1.

The size of sectors was different in individual shoots. FLARE-S expression in different leaf layers was also obvious. With the traditional selection for spectinomycin resistance, the transplastomic and wild-type sectors are not visible. Regeneration of plants with uniformly transformed plastid genomes was greatly facilitated by the fluorescing sectors expressing FLARE-S, which could be readily identified in UV light, dissected, and transferred for a second cycle of plant regeneration on spectinomycin-containing (500 mg/L) selective medium.

Given the high levels of FLARE-S accumulation we were interested to find out, if FLARE-S is toxic to plants. We expected that toxicity should be manifested as lower transformation efficiencies. Bombardment of 30 tobacco leaves with plasmids pMSK56 and pMSK57 yielded 71 and 89 spectinomycin resistant clones, respectively. Out of these, 61 and 77 lines were verified as transplastomic by fluorescence. Plastid transformation in a subset of these was confirmed by confocal laser scanning microscopy (7 clones each; see below) and Southern analysis (4 clones). The frequency of plastid transformation events with the FLARE-S -expressing genes

was slightly higher (~2 instead of ~1 per bombardment) than reported earlier with a chimeric *aadA* gene at the same insertion site (Svab and Maliga, 1993). Therefore, we assume that accumulation of FLARE-S at high levels is not detrimental. Lack of toxicity is also supported by the apparently normal phenotype of the plants in the greenhouse (not shown).

Localization of FLARE-S to tobacco plastids by confocal microscopy. Due to phenotypic masking, transplastomic and wild type sectors in a chimeric leaf are both green on a selective medium. However, we have found that in chimeric leaf sectors in the same cell some plastids express FLARE-S while others do not, when observed by confocal microscopy (Fig. 24). FLARE-S and chlorophyll fluorescence were detected separately in the fluorescein and rhodamine channels, respectively. The two images were then overlaid confirming that FLARE-S fluorescence derives from chloroplasts.

Expression of FLARE-S was also studied in non-green plastid types including the chromoplasts in petals and the non-green plastids in root cells (Fig. 24b,f). These studies were carried out in plants, which were homoplastomic for the transgenomes. Homoplastomic state was important, since in non-green tissues chlorophyll could not be used for confirmation of the organelles as plastids. Since FLARE-S expression could be readily detected in chloroplasts as well as non-green plastids, the plastid rRNA operon promoter is apparently active in all plastid types.

FLARE-S accumulation in tobacco leaves. Accumulation of FLARE-S in homoplastomic leaves was

tested using the commercially available GFP antibody, recognizing the GFP portion (239 amino acid residues) of FLARE16-S (520 amino acids). FLARE16-S1 (532 amino acids) was ~8 %, whereas FLARE16-S2 (532 amino acids) was ~18 % of total soluble leaf protein (Fig. 25). To calculate FLARE16-S concentrations, a GFP dilution series was used as a reference, and the values were than increased by 2.6 to correct for the larger size of the FLARE16-S1 and -S2 proteins.

Tracking plastid transformation in rice by FLARE-S expression. In rice, plant regeneration is from non-green embryogenic cells. Encouraged by FLARE-S expression in non-green tobacco plastids, we attempted to transform the non-green plastids of embryogenic rice tissue-culture cells. Plastid transformation was carried out using a rice-specific vector expressing FLARE11-S3 and targeting insertion of the *aadA11gfp-S3* gene in the *trnV/rps12/7* intergenic region. The location of the insertion site and the size of plastid targeting sequences in the rice vector are similar to the tobacco vectors shown in Fig. 23.

Plastid transformation in rice was carried out by bombardment of embryogenic rice suspension culture cells using gold particles coated with plasmid pMSK49 DNA. Rice cells, as most cereals, are naturally resistant to spectinomycin (Fromm *et al.*, 1987). FLARE-S, however, confers resistance to streptomycin as well (Svab and Maliga, 1993). Therefore, selection for transplastomic lines was carried out on selective streptomycin medium (100 mg/L). Streptomycin at this concentration inhibits the growth of embryogenic rice cells. After bombardment, the rice cells were first

selected in liquid embryogenic AA medium, then on the solid plant regeneration medium, on which the surviving resistant cells regenerated green shoots (12 in 25 bombarded plates). These shoots were rooted, and grown into plants. PCR amplification of border fragments in DNA isolated from the leaves of these plants confirmed integration of *aadA11gfp-S3* sequences in the plastid genome (Fig. 26). The left and right border fragments can not be amplified if the gene is integrated into the nuclear genome, as one of the primers (O4 or O6) of the pairs is outside the plastid targeting regions.

FLARE11-S3 expression in the leaves of two of the PCR-positive plants was tested by confocal laser-scanning microscopy. In rice, as in tobacco, the FLARE-S marker confirmed segregation of transplastomic and wild-type plastids (Fig. 27). In rice only a small fraction of chloroplasts expressed FLARE-S. Since individual cells marked with arrows in Fig. 27 contained a mixed population of wild-type and transgenic chloroplasts, FLARE-S in these cells could be expressed only from the plastid genome. Integration of *aadA11gfp-S3* into the nuclear genome downstream of plastid-targeting transit peptide would result in uniform expression of FLARE-S in each of the chloroplasts within the cell.

The sequences of the selectable marker genes of the invention are provided in Figures 28-34. Figure 35 depicts a table describing the selectable marker genes disclosed in the present example.

Direct visual identification of transplastomic sectors requires high level expression of FLARE-S in plastids. High GFP expression levels in Arabidopsis were toxic, interfering with plant regeneration. Toxicity of

wild-type (insoluble) GFP was linked to GFP accumulation in the nucleus and cytoplasm, and could be eliminated by targeting it to the endoplasmic reticulum (Haseloff *et al.*, 1997). GFP aggregates were also cytotoxic to *E. coli* cells (Crameri *et al.*, 1996). To enhance fluorescence intensity and to avoid cytotoxicity, soluble versions of the codon-modified GFP were obtained (Davis and Vierstra, 1998). We have utilized the gene for a soluble-modified GFP described by Davis and Vierstra (Davis and Vierstra, 1998) to create variants of FLARE-S, a fusion protein, which does not have an apparent cytotoxic effect. The frequency of plastid transformation, if affected at all, is increased rather then decreased. In tobacco, we normally obtain one transplastomic clone per bombarded leaf sample (Svab and Maliga, 1993), whereas with the FLARE-S genes on average we could recover two clones per sample. Plant regeneration from highly fluorescent tissue was readily obtained, and the regenerated plants have a phenotype indistinguishable from the wild type.

Plastid transformation in rice requires expression of the selective marker in non-green plastids. The rRNA operon has two promoters, one for the eubacterial-type (PEP) and one for the phage-type (NEP) plastid RNA polymerase. The promoter driving FLARE-S expression is recognized only by the eubacterial-type plastid RNA polymerase. Previously, it was assumed that the eubacterial-type promoter is active only in chloroplasts (Maliga, 1998). Accumulation of FLARE-S in roots and petals indicates that PEP is also active in non-green plastids.

Plastid transformation is a process that unavoidably yields chimeric plants, since cells of

higher plants contain a large number (300 to 50000) of plastid genome copies (Bendich, 1987), out of which initially only a few are transformed. High level expression of FLARE-S in plastids provides the means for visual identification of transplastomic sectors, even if they are present in a chimeric tissue. GFP and AAD could be expressed from two different genes in a plastid transformation vector. However, transformation with a marker gene encoding a bifunctional protein prevents separation of the two genes and simplifies engineering. The fluorescent selective marker will significantly reduce the work required to obtain genetically stable plastid transformants in tobacco, a species in which plastid transformation is routine. The bottleneck of applying plastid transformation in crop improvement is the lack of technology. In tobacco, chimeric clones with transformed plastids are readily identified by shoot regeneration (Svab *et al.*, 1990). In Arabidopsis, clones with transformed plastids are identified by greening (Sikdar *et al.*, 1998). We have shown here that FLARE-S is a suitable marker to select for transplastomes in embryogenic rice cells, which lack the visually identifiable tissue culture phenotypes exploited in tobacco and Arabidopsis. Data presented here are the first example for stable integration of foreign DNA into the rice plastid genome. These rice plants are heteroplastomic. Uniformly transformed rice plants will be obtained by further selection on streptomycin medium and screening the embryogenic cells for FLARE-S expression. Thus, the FLARE-S marker system will enable extension of plastid transformation to cereal crops.

The utility of the new chimeric promoters

The σ^{70}-type plastid ribosomal RNA operon promoter, Prrn, is the strongest known plastid promoter expressed in all tissue types. The ultimate product of this promoter in the plastid is RNA not protein. Therefore, a series of chimeric promoters were constructed to facilitate protein accumulation from Prrn, using expression of the neomycin phosphotransferase (NPTII) enzyme as the reference protein.

1) The expression cassettes have distinct tissue-specific expression profiles. Some of the expression cassettes described here will facilitate relatively high levels of protein expression in all tissues, including leaves, roots and seeds. Other cassettes have different expression profiles: for example will facilitate moderate levels of protein accumulation in the leaves while lead to relatively high levels of protein accumulation in the roots. Accumulation of a protein at levels of 10% to 50% of total soluble protein is considered high-level protein expression; low-levels of protein expression would be in the range of $\leq$0.1% total soluble cellular protein.

2) Efficiency of the selectable marker gene depends on the rate at which the gene product accumulates during the early stage of transformation. Since initially present only in a few copies per cell, high levels of expression from a few copies will provide protection from toxic substances early on, facilitating efficient recovery of transformed lines. The expression cassettes will be useful to drive the expression of the genes conferring resistance to the antibiotics

streptomycin, spectinomycin and hygromycin, and the herbicides phosphinotrycin and glyphosate. In such applications addition of amino acids at the N-terminus is acceptable, as long as it does not interfere with the expression of the selectable marker genes. NPTII is such an enzyme. In cases like NPTII, an N-terminal fusion and thereby the mRNA "Downstream Box" sequences give an additional at least two to four-fold increase in protein levels. The -DB construct which relied on an *NheI* site, and involved addition of one (N-terminal) amino acid of the source gene coding region is convenient, but is not necessary. When translational fusion is not feasible due to inactivation of proteins, seamless in-frame constructs may be created by PCR methods outlined in the application.

3) A second major area on which application of the chimeric promoters is extremely useful is protein expression for pharmaceutical, industrial or agronomic purposes. The examples include, but are not restricted to, production of vaccines, healthcare products like human hemoglobin, industrial or household enzymes.

REFERENCES

Allen, G.C., Hall, G.J., Michalowski, S., Newman, W., Spiker, S., Weissinger, A.K. and Thompson, W.F. (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. Plant Cell, 8, 899-913.

Allison LA, Simon LD, Maliga P (1996) Deletion of rpoB reveals a second distinct transcription system in plastids of higher plants. EMBO J 15: 2802-2809

Arntzen, C.J. (1997) High-tech herbal medicine: plant-based vaccines [news]. Nature Biotechnology, 15, 221-222.

Baulcombe, D. C. Chapman, S. and Cruz, S. S. 1995. Jellyfish green fluorescent protein as a reporter for virus infections. Plant J. 7: 1045-1053.

Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B. and Schaller, H. (1982) Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene, 19, 327-336.

Bendich, A. J. 1987. Why do chloroplasts and mitochondria contain so many copies of their genome? Bio-essays 6: 279-282.

Bonham-Smith, P.C. and Bourque, D.P. (1989) Translation of chloroplast-encoded mRNA: potential initiation and termination signals. Nucleic Acids Res. 17, 2057-2080.

Cao, J., Duan, X., McElroy, D., and Wu, R. (1992) Regeneration of herbicide resistant transgenic rice plants following microprojectile-mediated transformation of suspension culture cells. Plant Cell Report 11:586-591.

Carrer, H. and Maliga, P. (1995) Targeted insertion of foreign genes into the tobacco plastid genome without physical linkage to the selectable marker gene. Biotechnology, 13, 791-794.

Carrer, H., Hockenberry, T.N., Svab, Z. and Maliga, P. (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco. Mol. Gen. Genet., 241, 49-56.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. 1994. Green fluorescent protein as a marker for gene expression. Science 263: 802-805.

Chaudhuri, S. and Maliga, P. (1996) Sequences directing C to U editing of the plastid psbL mRNA are located within a 22 nucleotide segment spanning the editing site. EMBO J., 15, 5958-5964.

Chinault, A.C., Blakesley, V.A., Roessler, E., Willis, D.G., Smith, C.A., Cook, R.G., and Fenwick, R.G. 1986. Characterization of transferable plasmids for Shigella flexneri 2a that confer resistance trimethoprim, streptomycin and sulfonamides. Plasmid 15: 119-131.

Chiu, W-L., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. and Sheen, J. 1996. Engineered gfp as a vital reporter in plants. Curr. Biol. 6: 325-330.

Conrad, U. and Fiedler, U. (1998) Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. Plant Mol. Biol., 38, 101-109.

Corriveau, J.L., and Coleman, A.W. (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and the results for over 200 angiosperm species. Amer. J. Bot. 75:1443-1458.

Crameri, A., Whitehorn, E. A., Tate, E. and Stemmer, W. P. C. 1996. Improved green fluorescent protein by molecular evolution by DNA shuffling. Nature Biotech. 14: 315-319.

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A. and Tsien, R. Y. 1995. Understanding, improving and using green fluorescent proteins. Trends Biochem. Sci. 20: 448-455.

Dams, E., Hendriks, L., Van de Peer, Y., Neefs, J.M., Smits, G., Vandenbempt, I. and De Wachter, R. (1988) Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res., 16 Suppl, r87-173.

Daniell, H., Datta, R., Varma, S., Gray, S., and Lee, S.B. (1998) Containement of herbicide resistance through genetic engineering of the chloroplast genome. Nat. Biotech. 16:345-348.

Davis, S.J., and Vierstra, R.D. 1998. Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. Plant Mol. Biol. 36: 521-528.

De Block, M., Botterman, J., Vandewiele, M., Dockx, J., Thoen, C., Gossele, V., Rao, Movva, N., Thompson, C., Van Montagu, M., and Leemans, J. (1987). Engineering herbicide resistance in plants by expression of a detoxifying enzyme. EMBO J. 6:2513-2518.

De Block, M., De Brouwer, D., Tenning, P. (1989). Transformation of Brassica napus and Brassica oleracea using Agrobacterium tumefaciens and the expression of the bar and neo genes in the transgenic plants. Plant Physiol . 91:694-701.

Deana, A., Ehrlich, R. and Reiss, C. (1998) Silent mutations in the Escherichia coli ompA leader peptide region strongly affect transcription and translation in vivo. Nucleic Acids Res., 26, 4778-4782.

Ellis, R.J. (1979) The most abundant protein in the world. Trends Biochem. Sci., 4, 241-244.

Epel, B.L., Padgett, H.S., Heinlein, M., and Beachy, R. 1996. Plant virus movement protein dynamics probed with GFP-protein fusion. Gene 173: 75-79.

Etchegaray, J.P. and Inouye, M. (1999) Translational enhancement by an element downstream of the initiation codon in Escherichia coli. J. Biol. Chem., 274, 10079-10085.

Faxin, M., Plumbridge, J. and Isaksson, L.A. (1991) Codon choice and potential complementarity between mRNA

downstream of the initiation codon and bases 1471-1480 in 16S ribosomal RNA affects expression of glnS. Nucleic Acids Res., 19, 5247-5251.

Fromm, H., Edelman, M., Aviv, D. and Galun, E. 1987. The molecular basis ofrRNA-dependent spectinomycin resistance in Nicotiana chloroplasts. EMBO J. 11: 3233-3237.

Gallo-Meagher, M., and Irvine, J.E. (1996) Herbicide resistant transgenic sugarcane plants containing the bar gene. Crop Sci. 36:1367-1374.

Golds, T., Maliga, P., and Koop, H.U. 1993. Stable plastid transformation in PEG-treated protoplasts of Nicotiana tabaccum. Biotechnology 11: 95-97.

Gray, A.J., and Raybould, A.F. (1988) Reducing transgene escape routes. Nature 392:653-654.

Hajdukiewicz, P., Allison, L.A., Maliga, P. (1997) The two plastid RNA polymerases encoded by the nuclear and plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J. 16, 4041-4048.

Haseloff, J., Siemering, K. R., Prasher, D. C., and Hodge, S. 1997. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. Proc. Natl. Acad. Sci. USA 94: 2122-2127.

Hecker, K.H. and Roux, K.H. (1996) High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. Biotechniques, 20, 478-485.

Heim, R., Prasher, D. C. and Tsien, R. Y. 1994. Wavelength mutations and posttranslational autooxidation of green fluorescent protein. Proc. Natl. Acad. Sci. USA 91: 12501-12504.

Hibberd, J.M., Linley, P.J., Khan, M.S., and Gray, J.C. 1998. Transient expression of green fluorescent protein in various plastid types following micro-projectile bombardment. Plant J. 16: 627-632.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C-R., Meng, B-Y., Li, U-Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. 1989. The complete sequence of the rice (Oryza sativa) chloroplast genome: Intermolecular recombination between distict tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol. Gen. Genet. 217: 185-194.

Hirose, T. and Sugiura, M. (1996) Cis-acting elements and trans-acting factors for accurate translation of chloroplast psbA mRNAs: development of an

in vitro translation system from tobacco chloroplasts. EMBO J., 15, 1687-1695.

Houdt, H.V., Ingelbrecht, I., Montagu, M.V. and Depicker, A. (1997) Post-transcriptional silencing of a neomycin phosphotransferase II transgene correlates with the accumulation of unproductive RNAs and with increased cytosine methylation of 3' flanking positions. Plant J., 12, 379-392.

Ito, K., Kawakami, K. and Nakamura, Y. (1993) Multiple control of Escherichia coli lysyl-tRNA synthetase expression involves a transcriptional repressor and a translational enhancer element. Proc. Natl. Acad. Sci. USA, 90, 302-306.

Kane, J.F. (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Current Opinion In Biotechnology, 6, 494-500.

Khan, M.S. and Maliga, P. (1999) Fluorescent antibiotic resistance marker to track plastid transformation in higher plants. Nature Biotechnology, in press.

Kling, J. (1996) Could transgenic supercrops one day breed superweeds? Science 274:180-181.

Kvhler, R.H., Cao, J., Zipfel, W.R., Webb, W.W., and Hanson, M.R. 1997. Exchange of protein molecules through connections between higher plant plastids. Science 276: 2039-2042.

Kolodziej, P. A., and Young, R. A. 1991. Epitope tagging and protein surveillance. Methods Enzymol. 194: 508-519.

Koop, H.U., Steinm|ller, K., Wagner, H., Rvssler, C., Eibl, C., and Sacher, L. 1996. Integration of foreign sequences into the tobacco plastome via PEG-mediated protoplast transformation. Planta 199: 193-101.

Kusnadi, A., Nikolov, Z. and Howard, J. (1997) Production of recombinant proteins in transgenic plants: practical considerations. Biotechnology and Bioengineering, 56, 473-484.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature 227: 680-685.

Lefebvre, B., Formstecher, P. and Lefebvre, P. (1995) Improvement of the gene splicing overlap method (SOE) method. BioTechniques 19: 186-187

Maier, R.M., Neckermann, K., Igloi, G.L. and Kvssel, H. (1995) Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence

and fine tuning of genetic information by transcript editing. J. Mol. Biol., 251, 614-628.

Makrides, S.C. (1996) Strategies for achieving high-level expression of genes in Escherichia coli. Microbiological Reviews, 60, 512-538.

Maliga P (1995). Biolistic transformation of tobacco cells with nuclear drug resistance genes. In Maliga P, Klessig D, Cashmore A, Gruissem W, Varner J (eds), Methods in Plant Molecular Biology-A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 37-52.

Maliga, P. (1993) Towards plastid transformation in flowering plants. Trends in Biotechnology 11:101-107.

Maliga, P. 1998. Two plastid RNA polymerases of higher plants: an evolving story. Trends Plant Sci. 3: 4-6.

Mayfield, S.P., Cohen, A., Danon, A. and Yohn, C.B. (1994) Translation of the psbA mRNA of Chlamydomonas reinhardtii requires a structured RNA element contained within the 5' untranslated region. J. Cell Biol., 127, 1537-1545.

McBride, K.E., Scaaf, D.J., Daley, M. and Stalker, D. (1994) Controlled expression of plastid transgenes in plants based on a nuclear-encoded and plastid-targeted T7 RNA polymerase. Proc. Natl. Acad. Sci. USA 91, 7301-7305.

McBride, K.E., Svab, Z., Schaaf, D.J., Hogan, P.S., Stalker, D.M. and Maliga, P. (1995) Amplification of a chimeric Bacillus gene in chloroplasts leads to an extraordinary level of an insecticidal protein in tobacco. Biotechnology (N Y), 13, 362-365.

Medgyesy, P., Pay, A., and Marton, L. (1986) Transmission of paternal chloroplasts in Nicotiana. Mol. Gen. Genet. 204:195-198.

Mettler, I.J. 1987. A simple and rapid method for minipreparation of DNA from tissue-cultured plant cells. Plant Mol. Biol. Rep. 5: 346-349.

Mikkelsen, T.R., Andersen, B., and Jorgensen, R.B. (1996). The risk of crop transgene spread. Nature 380:31.

Misteli, T., and Spector, D.L. 1997. Applications of the green fluorescent protein in cell biology and biotechnology. Nature Biotech. 15: 961-964.

Mitta, M., Fang, L. and Inouye, M. (1997) Deletion analysis of cspA of Escherichia coli: requirement of the AT-rich UP element for cspA transcription and the downstream box in the coding region for its cold shock induction. Mol Microbiol, 26, 321-335.

Moll, B., Posby, L., and Maliga, P. 1990. Streptomycin and lincomycin resistance are selective plastid markers in cultured Nicotiana cells. Mol. Gen. Genet. 221: 245-250.

Muller, A.J., and Grafe, R. 1978. Isolation and characterization of cell lines of Nicotiana tobacum lacking nitrate reductase. Mol. Gen. Genet. 161: 67-76.

Murashige, T. and Skoog, F. (1962) A revised medium for the growth and bioassay with tobacco tissue culture. Physiologia Plantarum, 15, 473-497.

Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M. and Rochaix, J.D. (1999) Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in Chlamydomonas. The Plant Cell, 11, 957-970.

Olins, P.O., Devine, C.S., Rangwala, S.H., Kavka, K.S. (1988) The T7 phage gene 10 leader RNA, a ribosome-binding site that dramatically enhances the expression of foreign genes in Escherichia coli. Gene 73: 227-235.

O'Neill, C., Horvath, G.V., Horvath, E., Dix, P.J., and Medgyesy, P. 1993. Chloroplast transformation in plants: polyethylene glycol (PEG) treatment of protoplasts is an alternative to biolistic delivery system. Plant J. 3: 729-738.

Orozco,E.M., Chen,L.J. and Eilers,R.J. (1990) The divergently transcribed rbcL and atpB genes of tobacco plastid DNA are separated by nineteen base pairs. Curr. Genet., 17, 65-71.

Pang, S-Z., DeBoer, D.L., Wan, Y., Ye, G., Layton, J.G., Neher, M.K., Armstrong, C.L., Fry, J.E., Hinchee, M.A.W., and Fromm, M.E. 1996. An improved green fluorescent protein gene as avital marker in plants. Plant Physiol. 112: 893-900.

Prasher, D. C. 1995. Using GFP to see the light. Trends Genet. 11: 320-323.

Prasher, D. C., Eckenrode, V. K., Ward, W. W., Predergast, F. G. and Cormier, M. J. 1992. Primary structure of the Aequorea victoria green-fluorescent protein. Gene 111: 229-233.

Reichel, C., Mathur, J., Eckes, P., Langenkemper, K., Koncz, C., Schell, J., Reiss, B., and Maas, C. 1996. Enhanced green fluorescence by the expression of an Aequorea victoria green fluorescent protein mutant in mono- and dicotyledonous plant cells. Proc. Natl. Acad. Sci. USA 93: 5888-5893.

Rochaix, J.D. (1996) Post-transcriptional regulation of chloroplast gene expression in

Chlamydomonas reinhardtii. Plant Mol. Biol., 32, 327-341.

Rouwendal., G.J.A., Mendes, O., Wolbert, E.J.H., and de Boer, A.D., 1997. Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage. Plant Mol. Biol. 33: 989-999.

Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.A., and Allard, R.W. (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA 81: 8014-8018.

Sakamoto, W., Kindle, K.L. and Stern, D.B. (1993) In vivo analysis of Chlamydomonas chloroplast petD gene expression using stable transformation of beta-glucuronidase translational fusions. Proc. Natl. Acad. Sci. USA, 90, 497-501.

Sawasaki, T., Seki, M., Anzai, H., Irifune, K., and Morikawa, H. (1994). Stable transformation of Arabidopsis with the bar gene using particle bombardment. Transgenic Res. 3:279-286.

Serino, G., and Maliga, P. (1997) A negative selection scheme based on the expression of cytosine deaminase in plastids. Plant J . 12:697-701.

Sheen, J., Hwang, S., Niwa, Y., Kobayashi, H., and Galbraith, D.W. 1995. Green fluorescent protein as a new vital marker in plant cells. Plant J. 8: 777-784.

Shimada, H., and Sugiura, M. (1991) Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes. Nucleic Acids Res. 19:983-995.

Shinozaki, K. and Sugiura, M. (1982) The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose-1,5-bisphosphate carboxylase / oxygenase. Gene 20: 91-102.

Shinozaki, K., Deno, H., Wakasugi, T. and Sugiura, M. (1986a) Tobacco chloroplast gene coding for subunit I of proton-translocating ATPase: comparison with the wheat subunit I and E. coli subunit b. Curr. Genet., 10, 421-423.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsabayashi, T., Zaita, N., Chungwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Deno, H., Kamogashira, T., Yamada, K., Kasuda, J., Takaiwa, F., Kato, A., Todoh, N., Shimada, H. and Sugiura, M. (1986b) The complete sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J., 5, 2043-2049.

Sikdar, S.R., Serino, G., Chaudhuri, S., and Maliga, P. 1998. Plastid transformation in Arabidopsis thaliana. Plant Cell Rep. 18: 20-24.

Spencer, T.M., Gordon-Kamm, W.J., Daines, R.J., Start, W.G., and Lemaux, P.G. (1990). Bialaphos selection of stable transformants from maize cell culture. Theor. Appl. Genet. 79:625-631.

Sprengart, M.L., Fuchs, E. and Porter, A.G. (1996) The downstream box: an efficient and independent translation initiation signal in Escherichia coli. EMBO J., 15, 665-674.

Sriraman, P., Silhavy, D. and Maliga, P. (1998a) The phage-type PclpP-53 plastid promoter comprises sequences downstream of the transcription initiation site. Nucleic Acids Res., 26, 4874-4879.

Sriraman, P., Silhavy, D. and Maliga, P. (1998b) Transcription from heterologous rRNA operon promoters in chloroplasts reveals requirement for specific activating factors. Plant Physiol., 117, 1495-1499.

Sriraman, P., Silhavy, D., Maliga, P. (1998b) Transcription from heterologous ribosomal RNA operon promoters in chloroplasts reveals requirement for specific activating factors. Plant Physiol. 117: 1495-1499.

Staub, J., Maliga, P. (1993) Accumulation of D1 polypeptide in tobacco plastids is regulated via the untranslated region of the psbA mRNA. EMBO J. 12:601-606

Staub, J.M. and Maliga, P. (1994) Translation of psbA mRNA is regulated by light via the 5'-untranslated region in tobacco plastids. Plant J., 6, 547-553.

Stemmer, W.P., Crameri, A., Ha, K.D., Brennan, T.M. and Heyneker, H.L. (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene, 164, 49-53.

Stern, D.B., Higgs, D.C. And Yang, J. (1997) Transcription and translation in chloroplasts. Trends Plant Sci. 2, 308-315.

Stiekema WJ, Heidekamp F, Dirkse WG, van Beckum J, deHaan P, ten Bosch C, Louwerse JD (1988) Molecular cloning and analysis of four potato tuber mRNAs. Plant Mol Biol 11: 255-269

Strauch, E., Wohlleben, , W., and P|hler, A. (1988) Cloning of phosphinothricin acetyltransferase gene from Streptomyces viridochromogenes T|4494 and its expression in Streptomyces lividans and Escherichia coli.. Gene 63:65-74.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to

direct expression of cloned genes. Methods Enzymol. 185: 60-89.

Sugita, M. and Sugiura, M. (1984) Nucleotide sequence and transcription of the gene for the 32,000 dalton thylakoid membrane protein from Nicotiana tabacum. Mol. Gen. Genet. 195: 308-313.

Svab, Z., Hajdukiewicz, P., and Maliga, P. 1990. Stable transformation of plastids in higher plants. Proc. Natl. Acad. Sci. USA 87: 8526-8530.

Svab, Z., and Maliga, P. (1993). High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. Proc. Natl. Acad. Sci. USA 90:913-917.

Tachibana. K., Watanabe. T., Sekizawa, T., and Takematsu, T. (1986) Action mechanism of bialaphos II Accumulation of ammonia in plants treated with bialaphos. J. Pest. Sci. 11:33-37.

Tanaka, M., Obokata, J., Chunwongse, J., Shinozaki, K., Sugiura, M. (1987) Rapid splicing and stepwise processing of a transcript from the psbB operon in tobacco chloroplasts: Determination of the intron sites in petB and petD. Mol. Gen. Genet. 209: 427-431.

Thompson, C.J., Movva, N.R,, Tizard, R., Crameri, R., Davies, J.E., Lauwereys, M., Botterman, J. (1987) Characterization of the herbicide-resistance gene bar from Streptomyces hygroscopicus. EMBO J. 6:2519-2523.

Timmons, A.M., Charters, Y.M., Crawford, J.W., Burn, D., Scott, S.E., Subbels, S.J., Wilson, N.J., Robertson, A., O'Brian, E.T., Squire, G.R., and Wilkinson, M.J. (1996) Risks from transgenic crops. Nature 380:487.

Tompson, J.A., Abdullah, R., and Cocking, E.C. 1986. Protoplast culture of rice using media solidified with agarose. Plant Science 47: 123-133.

Uchida, K (1992) Recombination and amplification of multiple portions of genomic DNA by a modified polymerase chain reaction. Anal. Biochem. 202: 159-161.

Varshavsky, A. (1996) The N-end rule: Functions, mysteries, uses. Proc. Natl. Acad. Sci USA 93, 121423-12149.

Vera, A., and Sugiura M (1995). Chloroplast rRNA transcription from structurally different tandem promoters: an additional novel-type promoter. Curr. Genet. 27, 280-284.

Voorma, H.O. (1996) Control of translation in prokaryotes. In: Translational Control, Hershey, J.W.B., Methews, M.B. and Sonenberg, N., eds. pp. 759-777, Cold Spring Harbor Laboratory Press, Plainview, NY

Wehrmann, A., Vliet, V.A., Opsomer, C., Bottermanm J., and Schulz, A. (1996) The bar and pat gene products make them equally applicable for plant engineers. Nat. Biotechnol. 14:1274-1278.

Wohlleben, W., Arnold, W., Broer, I., Hillemann, D., Strauch, E. and P|hler, A. (1988) Nucleotide sequence of the phosphinothricin N-acetyltransferase gene from Streptomyces T|949 and its expression in Nicotiana tabacum. Gene 70:25-37.

Wu, C.J. and Janssen, G.R. (1996) Translation of vph mRNA in Streptomyces lividans and Escherichia coli after removal of the 5' untranslated leader. Mol Microbiol, 22, 339-355.

Zerges, W., Girard-Bascou, J. and Rochaix, J.D. (1997) Translation of the chloroplast psbC mRNA is controlled by interactions between its 5' leader and the nuclear loci TBC1 and TBC3 in Chlamydomonas reinhardtii. Mol. Cell. Biol., 17, 3440-3448.

Zhang, W., and Wu, R. 1988. Efficient regeneration of transgenic plants from rice protoplasts and correctly regulated expression of the foreign gene in the plants. Theor. Appl. Gene. 76: 835-840.

Zhou, J., Liu, W.J., Peng, S.W., Sun, X.Y. and Frazer, I. (1999) Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. Journal Of Virology, 73, 4972-4982.

Zoubenko, O. V., Allison, L. A., Svab, Z. and Maliga, P. 1994. Efficient targeting of foreign genes into the tobacco plastid genome. Nuceic Acids Res. 22: 3819-3824.

While certain of the preferred embodiments of the present invention have been described and specifically exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made thereto without departing from the scope and spirit of the present invention, as set forth in the following claims.

What is claimed is:

1. A recombinant DNA construct for expressing at least one heterologous protein in the plastids of higher plants, said construct comprising a 5' regulatory region which includes a promoter element, a leader sequence and a downstream box element operably linked to a coding region of said at least one heterologous protein, said chimeric regulatory region enhancing translational efficiency of an mRNA molecule encoded by said DNA construct.

2. A vector comprising the DNA construct of claim 1.

3. A recombinant DNA construct as claimed in claim 1, said 5' regulatory region being selected from the group consisting of PrnnLatpB+DBwt, SEQ ID NO:1, PrrnLatpB-DB, SEQ ID NO:2, PrrnLatpB+DBm, SEQ ID NO:3, PrrnLclpP+DBwt, SEQ ID NO: 4, PrrnclpP-DB, SEQ ID NO:5, PrrnLrbcL+DBwt, SEQ ID NO:6, PrrnLrbcL-DB, SEQ ID NO:7, PrrnLrbcL+DBm, SEQ ID NO:8, PrrnLpsbB+DBwt, SEQ ID NO:9, PrrnLpsbB-DB, SEQ ID NO:10, PrrnLpsbA+DBwt, SEQ ID NO: 11, PrrnLpsbA-DB, SEQ ID NO:12, PrrnLpsbA-DB(+GC), SEQ ID NO:13.

4. A recombinant DNA construct as claimed in claim 1, said 5' regulatory region being selected from the group consisting of PrrnLT7g10+DB/Ec, SEQ ID NO:14,

PrrnLT7g10+DB/pt, SEQ ID NO:15, PrrnLT7g10-DB, SEQ ID NO:15.

5. A vector comprising a DNA construct as claimed in claim 1.

6. A DNA construct as claimed in claim 1, said downstream box element having a sequence selected from the group consisting of 5'TCCAGTCACTAGCCCTGCCTTCGGCA'3 and 5'CCCAGTCATGAATCACAAAGTGGTAA'3.

7. A DNA construct as claimed in claim 1, wherein said heterologous protein is expressed from a *bar* gene encoded by *S. hydroscopicus* said bar gene inserted into a plasmid selected from the group consisting of pKO12, and pJEK3, said pJEK3 having the sequence of SEQ ID NO: 18.

8. A DNA construct as claimed in claim 1, wherein said heterologous protein is expressed from a synthetic bar encoding nucleic acid, said synthetic bar nucleic acid having selected from the group consisting of SEQ ID NO: 19 and SEQ ID NO:20.

9. A DNA construct as claimed in claim 1, said at least one heterologous protein comprising a fusion protein.

10. A DNA construct as claimed in claim 9, said fusion protein having a first and second coding region operably linked to said 5' regulatory region such that production of said fusion protein is regulated by

said 5' regulatory region, said first coding region encoding a selectable marker gene and said second coding region encoding a fluorescent molecule to facilitate visualization of transformed plant cells.

11. A vector comprising the DNA construct of claim 10.

12. A DNA construct as claimed in claim 9, said fusion protein consisting of an aadA coding region operably linked to a green fluorescent protein coding region.

13. A DNA construct as claimed in claim 10, said aadA coding region being operably linked to said green fluorescent protein coding region via a nucleic acid molecule encoding a peptide linker having a sequence selected from the group consisting of ELVEGKLELVEGLKVA and ELAVEGKLEVA.

14. A DNA construct as claimed in claim 10, said construct having a sequence selected from the group of SEQ ID NOS: 21-25 and 27.

15. A plasmid for transforming the plastids of higher plants, said plasmid being selected from the group consisting of pHK30(B), pHK31(B), pHK60, pHK32(B), pHK33(B), pHK34(A), pHK35(A), pHK64(A), pHK36(A), pHK37(A), pHK38(A), pHK39(A), pHK40(A), pHK41(A), pHK42(A), pHK43(A), pMSK56, pMSK57, pMSK48, pMSK49, pMSK35, pMSK53 and pMSK54.

16. A transgenic plant containing a plasmid as claimed in claim 15.

17. A transgenic plant as claimed in claim 15, said plant being selected from the group consisting of monocots and dicots.

18. A method for producing transplastomic monocots, comprising:

a) obtaining embryogenic cells;

b) exposing said cells to a heterologous DNA molecule under conditions whereby said DNA enters the plastids of said cells, said heterologous DNA molecule encoding at least one exogenous protein, said at least one exogenous protein encoding a selectable marker;

c) applying a selection agent to said cells to facilitate sorting of untransformed plastids from transformed plastids, said cells containing transformed plastids surviving and dividing in the presence of said selection agent;

d) transferring said surviving cells to selective media to promote shoot regeneration and growth; and

e) rooting said shoots, thereby producing transplastomic monocot plants.

19. A method as claimed in claim 18, wherein said heterologous DNA molecule is introduced into said plant cell via a process selected from the group consisting of biolistic bombardment, Agrobacterium- mediated transformation, microinjection and electroporation.

20. A method as claimed in claim 18, wherein protoplasts are obtained from said embryogenic cells and said heterologous DNA molecule is delivered to said protoplasts by exposure to polyethylene glycol.

21. A method as claimed in claim 18, wherein said selection agent is selected from the group consisting of streptomycin, and paromomycin

22. A monocot transformed via the method of claim 18.

23. A transformed monocot plant as claimed in claim 22, said monocot plant being selected from the group consisting of maize, millet, sorghum, sugar cane, rice, wheat, barley, oat, rye, and turf grass.

24. A method for producing transplastomic rice plants, said method comprising:

a) obtaining embryogenic calli;

b) inducing proliferation of calli on modified CIM medium;

c) obtaining embryogenic cell suspensions of said proliferating calli in liquid AA medium;

d) bombarding said embryogenic cells with microprojectiles coated with plasmid DNA;

e) tranferring said bombarded cells to selective liquid AA medium;

f) transferring said cells surviving in AA medium to selective RRM regeneration medium for a time period sufficient for green shoots to appear; and

g) rooting said shoots in a selective MS salt medium.

25. A method as claimed in claim 24, said plasmid DNA being selected from the group of plasmids consisting of pMSK35 and pMSK53, pMSK54 and pMSK49.

26. A transplastomic rice plant produced by the method of claim 24.

27. A method for containing transgenes in transformed plants, comprising:

a) determining the codon usage in said plant to be transformed and in microbes found in association with said plant; and

b) genetically engineering said transgene sequence via the introduction of rare codons to abrogate expression of said transgene in said plant associated microbe.

28. A method as claimed in claim 27, wherein said transgene is a bar gene and said rare codons are arginine encoding codons selected from the group consisting of AGA and AGG.

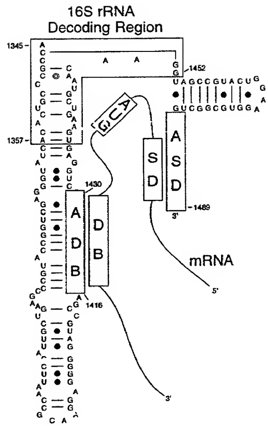

16S rRNA
Decoding Region
1345
1452
1357
1430
1489
1416
ADB
DB
SD
ASD
GCA
mRNA
5'
3'
3'

Figure 1A

	1 10 20 26	
pt ADB	3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5'	
	1430 1416	
	1 10 20 26	
Ec ADB	3'-GGGUCAGUACUUAGUGUUUCACCAUU-5'	
	1483 1469	

Figure 1B

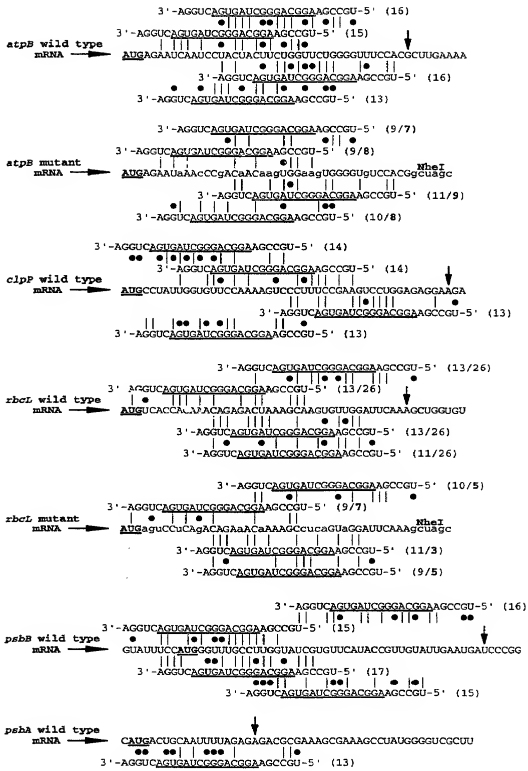

3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (16)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (15)
atpB wild type
mRNA
AUGAGAAUCAAUCCUACUACUUCUGGUUCUGGGGUUUCCACGCUUGAAAA
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (16)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (9/7)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (9/8)
atpB mutant
mRNA
NheI
AUGAGAAUaAAcCCgACaACaagUGGaagUGGGGUgUCCACGgcuagc
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (11/9)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (10/8)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (14)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (14)
clpP wild type
mRNA
AUGCCUAUUGGUGUUCCAAAAGUCCCUUUCCGAAGUCCUGGAGAGGAAGA
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13/26)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13/26)
rbcL wild type
mRNA
AUGUCACCACAAACAGAGACUAAAGCAAGUGUUGGAUUCAAAGCUGGUGU
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13/26)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (11/26)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (10/5)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (9/7)
rbcL mutant
mRNA
NheI
AUGaguCCuCAgACAGAaACaAAAGCcucaGUaGGAUUCAAAgcuagc
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (11/3)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (9/5)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (16)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (15)
psbB wild type
mRNA
GUAUUUCCAUGGGUUUGCCUUGGUAUCGUGUUCAUACCGUUGUAUUGAAUGAUCCCGG
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (17)
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (15)
psbA wild type
mRNA
CAUGACUGCAAUUUUAGAGAGACGCGAAAGCGAAAGCCUAUGGGGUCGCUU
3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13)

Figure 2A

```
T7g10 mRNA          AUGGCUAGCAUGACUGGUGGACAGCAAAUGGGUCGCGGAUCCGGCUGCUA
                         |   ||||| |●●●● |  | ||
            Ec ADB 3'-GGGUCAGUACUUAGUGUUUCACCAUU-5' (15)

                                             NheI
T7g10+DB/Ec mRNA    AUGGCaAGCAUGACUGGUGGACAGgcuagc
                         ||  ||● | ●  ||   || ●|
            pt ADB 3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (13)

                                             NheI
T7g10+DB/pt mRNA    AUGGCaAucacuagcccugccuuGgcuagc
                        || |||||||||||||||| | ●|
            pt ADB 3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (21)

                             NheI         neo
T7g10-DB mRNA        ACAUAUGgcuagcauugaacaagauggauugcau
                      |   | ●||||||●||     | ●●
            pt ADB 3'-AGGUCAGUGAUCGGGACGGAAGCCGU-5' (14)
```

Figure 2B

09/762105

PrrnLatpB+DBwt (pHK10)

```
                SacI
    1  gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
   51  GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG AATTAACCGA
  101  TCGACGTGCa AGCGGACATT TATTTTaAAT TCGATAATTT TTGCAAAAAC
  151  ATTTCGACAT ATTTATTTAT TTTATTATTA TGAGAATCAA TCCTACTACT
                                     NheI
  201  TCTGGTTCTG GGGTTTCCAC Ggctagc
```

PrrnLatpB-DB (pHK11)

```
                SacI
    1  gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
   51  GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG AATTAACCGA
  101  TCGACGTGCa AGCGGACATT TATTTTaAAT TCGATAATTT TTGCAAAAAC
                                                NheI
  151  ATTTCGACAT ATTTATTTAT TTTATTATTA TGAGAgctag c
```

PrrnLatpB+DBm (pHK50)

```
                SacI
    1  gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
   51  GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG AATTAACCGA
  101  TCGACGTGCa AGCGGACATT TATTTTaAAT TCGATAATTT TTGCAAAAAC
  151  ATTTCGACAT ATTTATTTAT TTTATTATTA TGAGAATaAA cCCgACaACa
                                     NheI
  201  agTGGaagTG GGGTgTCCAC Ggctagc
```

PrrnLclpP+DBwt (pHK12)

```
                SacI
    1  gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
   51  GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG TTACGTTTCC
  101  ACCTCAAAGT GAAATATAGT ATTTAGTTCT TTCTTTCATT TAATGCCTAT
                                                NheI
  151  TGGTGTTCCA AAAGTCCCTT TCCGAAGTCC TGGAGAGGAA gctagc
```

PrrnLclpP-DB (pHK13)

```
                SacI
    1  gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
   51  GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG TTACGTTTCC
                                                          NheI
  101  ACCTCAAAGT GAAATATAGT ATTTAGTTCT TTCTTTCATT TAATGCCTgc
  151  tagc
```

Figure 3A

PrrnLrbcL+DBwt (pHK14)

```
         SacI
    1    gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

   51    GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG TCGAGTAGAC

  101    CTTGTTGTTG TGAaAATTCT TAATTCATGA GTTGTAGGGA GGGATTTATG
                                                    NheI
  151    TCACCACAAA CAGAGACTAA AGCAAGTGTT GGATTCAAAg ctagc
```

PrrnLrbcL-DB (pHK15)

```
         SacI
    1    gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

   51    GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG TCGAGTAGAC

  101    CTTGTTGTTG TGAaAATTCT TAATTCATGA GTTGTAGGGA GGGATTTATG
            NheI
  151    TCAgctagc
```

PrrnLrbcL+DBm (pHK54)

```
         SacI
    1    gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

   51    GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG TCGAGTAGAC

  101    CTTGTTGTTG TGAaAATTCT TAATTCATGA GTTGTAGGGA GGGATTTATG
                                                    NheI
  151    aguCCuCAgA CAGAaACaAA AGCcucaGTa GGATTCAAAg ctagc
```

PrrnLpsbB+DBwt (pHK16)

```
         SacI
    1    gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

   51    GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG CAATGCAATA

  101    AAGTTACGTA GTGTCTATTT ATCTTTGATA TAAGGGGTAT TTCCATGGGT
                                                    NheI
  151    TTGCCTTGGT ATCGTGTTCA TACCGTTGTA TTGAATGATg ctagc
```

PrrnLpsbB-DB (pHK17)

```
         SacI
    1    gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

   51    GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG CAATGCAATA
                                                      NcoI NheI
  101    AAGTTACGTA GTGTCTATTT ATCTTTGATA TAAGGGGTAT TTccatggct

  151    agc
```

Figure 3B

PrrnLpsbA+DBwt (pHK21)

```
          SacI
  1   gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

 51   GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAA AAAAGCCTTC

101   CATTTTCTAT TTTGATTTGT AGAAAACTAG TGTGCTTGGG AGTCCCTGAT
                                                     NheI
151   GATTAAATAA ACCAAGATTT TACCATGACT GCAATTTTAG AGAGAgctag

201   c
```

PrrnLpsbA-DB (pHK22)

```
          SacI
  1   gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

 51   GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAA AAAAGCCTTC

101   CATTTTCTAT TTTGATTTGT AGAAAACTAG TGTGCTTGGG AGTCCCTGAT
                                 NcoI NheI
151   GATTAAATAA ACCAAGATTT TAccatggct agc
```

PrrnLpsbA-DB(+GC) (pHK23)

```
          SacI
  1   gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG

 51   GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG CAAAAAGCCT

101   TCCATTTTCT ATTTTGATTT GTAGAAAACT AGTGTGCTTG GGAGTCCCTG
                                   NcoI NheI
151   ATGATTAAAT AAACCAAGAT TTTAccatgg ctagc
```

Figure 3C

PrrnLT7g10+DB/Ec (pHK18)

```
         SacI
  1      gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
                                                  ●
 51      GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG GGAGACCACA

101      ACGGTTTCCC aCTAGAAATA ATTTTGTTTA ACTTTAAGAA GGAGATATAC
                                NheI
151      ATATGGCaAG CATGACTGGT GGACAGgcta gc
```

PrrnLT7g10+DB/pt (pHK19)

```
         SacI
  1      gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
                                                  ●
 51      GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG GGAGACCACA

101      ACGGTTTCCC aCTAGAAATA ATTTTGTTTA ACTTTAAGAA GGAGATATAC
                                NheI
151      ATATGGCaAt cactagccct gccttGgcta gc
```

PrrnLT7g10-DB (pHK20)

```
         SacI
  1      gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
                                                  ●
 51      GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG GGAGACCACA

101      ACGGTTTCCC aCTAGAAATA ATTTTGTTTA ACTTTAAGAA GGAGATATAC
             NheI
151      ATATGgctag c
```

Figure 3D

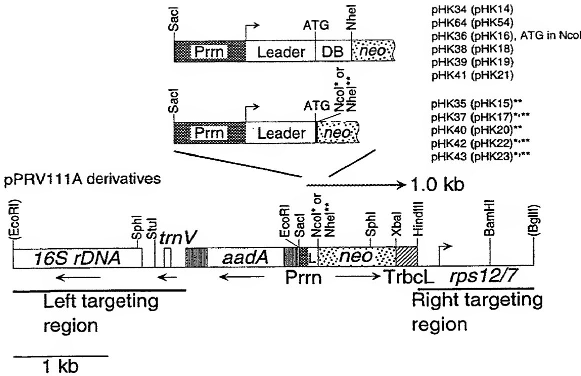

SacI
ATG
NheI
Prrn
Leader
DB
neo
pHK34 (pHK14)
pHK64 (pHK54)
pHK36 (pHK16), ATG in NcoI
pHK38 (pHK18)
pHK39 (pHK19)
pHK41 (pHK21)
SacI
ATG
NcoI* or NheI**
Prrn
Leader
neo
pHK35 (pHK15)**
pHK37 (pHK17)*,**
pHK40 (pHK20)**
pHK42 (pHK22)*,**
pHK43 (pHK23)*,**
pPRV111A derivatives
1.0 kb
(EcoRI)
SphI
StuI
trnV
EcoRI
SacI
NcoI* or NheI**
SphI
XbaI
HindIII
BamHI
(BglII)
16S rDNA
aadA
L
neo
Prrn
TrbcL
rps12/7
Left targeting region
Right targeting region
1 kb

Figure 4A

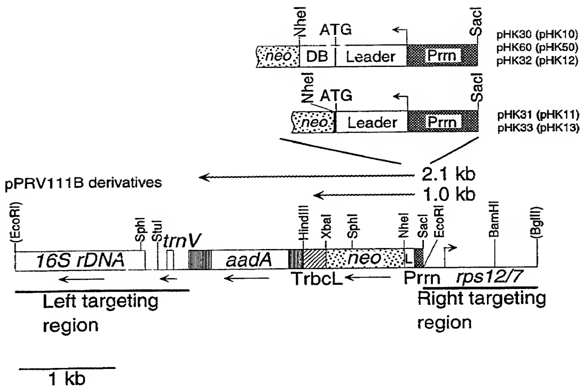

NheI
ATG
SacI
neo
DB
Leader
Prrn
pHK30 (pHK10)
pHK60 (pHK50)
pHK32 (pHK12)
NheI
ATG
SacI
neo
Leader
Prrn
pHK31 (pHK11)
pHK33 (pHK13)
2.1 kb
1.0 kb
pPRV111B derivatives
(EcoRI)
SphI
StuI
trnV
HindIII
XbaI
SphI
NheI
SacI
EcoRI
BamHI
(BglII)
16S rDNA
aadA
neo
L
TrbcL
Prrn
rps12/7
Left targeting region
Right targeting region
1 kb

Figure 4B

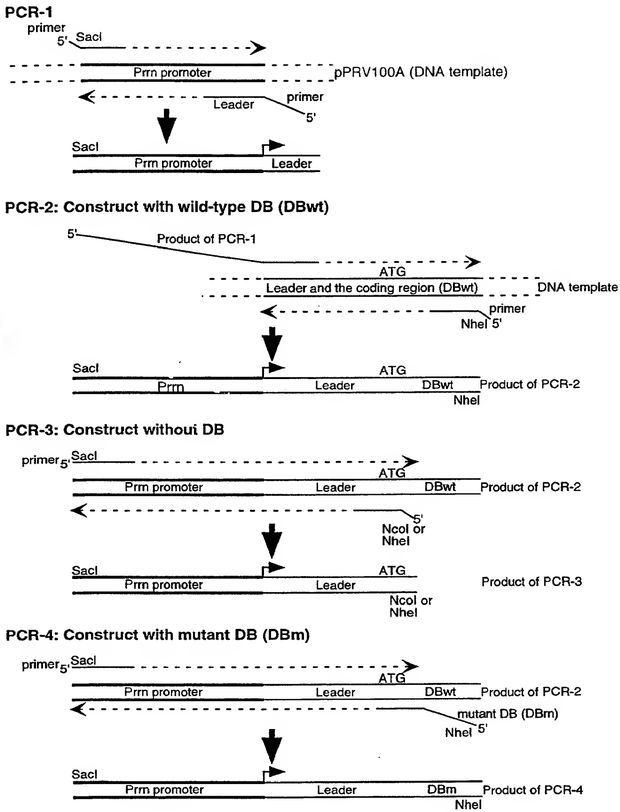

PCR-1
primer
5'
SacI
Prrn promoter
pPRV100A (DNA template)
Leader
primer
5'
SacI
Prrn promoter
Leader
PCR-2: Construct with wild-type DB (DBwt)
5'
Product of PCR-1
ATG
Leader and the coding region (DBwt)
DNA template
primer
NheI 5'
SacI
ATG
Prrn
Leader
DBwt
Product of PCR-2
NheI
PCR-3: Construct without DB
primer 5' SacI
ATG
Prrn promoter
Leader
DBwt
Product of PCR-2
5'
NcoI or NheI
SacI
ATG
Prrn promoter
Leader
Product of PCR-3
NcoI or NheI
PCR-4: Construct with mutant DB (DBm)
primer 5' SacI
ATG
Prrn promoter
Leader
DBwt
Product of PCR-2
mutant DB (DBm)
NheI 5'
SacI
Prrn promoter
Leader
DBm
Product of PCR-4
NheI

Figure 5

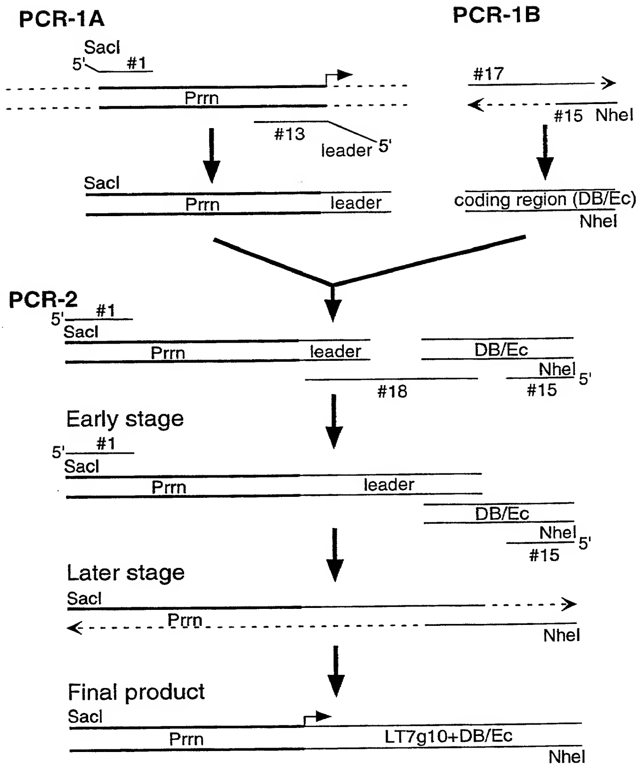

PCR-1A
SacI
5'
#1
Prrn
#13
leader
5'
SacI
Prrn
leader
PCR-1B
#17
#15
NheI
coding region (DB/Ec)
NheI
PCR-2
5'
#1
SacI
Prrn
leader
DB/Ec
NheI
5'
#18
#15
Early stage
5'
#1
SacI
Prrn
leader
DB/Ec
NheI
5'
#15
Later stage
SacI
Prrn
NheI
Final product
SacI
Prrn
LT7g10+DB/Ec
NheI

Figure 6

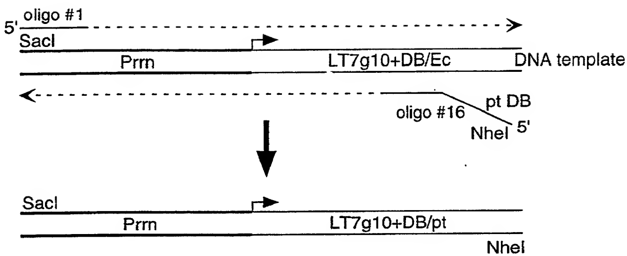

5'
oligo #1
SacI
Prrn
LT7g10+DB/Ec
DNA template
pt DB
oligo #16
NheI
5'
SacI
Prrn
LT7g10+DB/pt
NheI

Figure 7

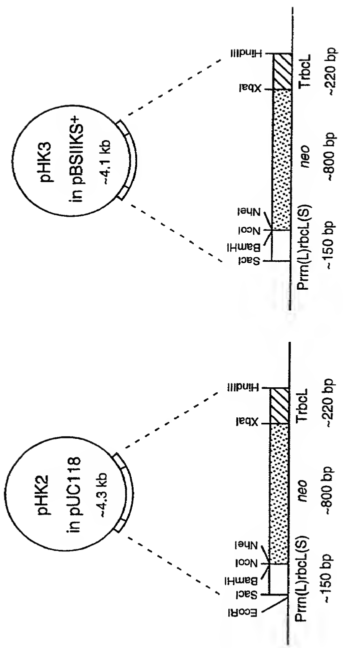

pHK3
in pBSIIKS+
~4.1 kb
SacI
BamHI
NcoI
NheI
XbaI
HindIII
Prrn(L)rbcL(S)
neo
TrbcL
~150 bp
~800 bp
~220 bp
pHK2
in pUC118
~4.3 kb
EcoRI
SacI
BamHI
NcoI
NheI
XbaI
HindIII
Prrn(L)rbcL(S)
neo
TrbcL
~150 bp
~800 bp
~220 bp

Figure 8

```
     SacI
   1 gagctcggta cccaaaGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG

  51 AGGCTCGTGG GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG
                                                          NcoI
 101 CGAACTCCGG GCGAATAcGA AGCGCtTGGA TACAGTTGTA GGGAGGGATc
          NheI
 151 catggctagc ATTGAACAAG ATGGATTGCA CGCAGGTTCT CCGGCCGCTT

 201 GGGTGGAGAG GCTATTCGGC TATGACTGGG CACAACAGAC AATCGGCTGC

 251 TCTGATGCCG CCGTGTTCCG GCTGTCAGCG CAGGGGCGCC CGGTTCTTTT

 301 TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAACTCCAG GACGAGGCAG

 351 CGCGGCTATC GTGGCTGGCC ACGACGGGCG TTCCTTGCGC AGCTGTGCTC

 401 GACGTTGTCA CTGAAGCGGG AAGGGACTGG CTGCTATTGG GCGAAGTGCC

 451 GGGGCAGGAT CTCCTGTCAT CTCACCTTGC TCCTGCCGAG AAAGTATCCA

 501 TCATGGCTGA TGCAATGCGG CGGCTGCATA CGCTTGATCC GGCTACCTGC

 551 CCATTCGACC ACCAAGCGAA ACATCGCATC GAGCGAGCAC GTACTCGGAT

 601 GGAAGCCGGT CTTGTCGATC AGGATGATCT GGACGAAGAG CATCAGGGGC

 651 TCGCGCCAGC CGAACTGTTC GCCAGGCTCA AGGCGCGCAT GCCCGACGGC

 701 GAGGATCTCG TCGTGACACA TGGCGATGCC TGCTTGCCGA ATATCATGGT

 751 GGAAAATGGC CGCTTTTCTG GATTCATCGA CTGTGGCCGG CTGGGTGTGG

 801 CGGACCGCTA TCAGGACATA GCGTTGGCTA CCCGTGATAT TGCTGAAGAG

 851 CTTGGCGGCG AATGGGCTGA CCGCTTCCTC GTGCTTTACG GTATCGCCGC

 901 TCCCGATTCG CAGCGCATCG CCTTCTATCG CCTTCTTGAC GAGTTCTTCT
                  XbaI
 951 GAgcgggtct agagtAGACA TTAGCAGATA AATTAGCAGG AAATAAAGAA

1001 GGATAAGGAG AAAGAACTCA AGTAATTATC CTTCGTTCTC TTAATTGAAT

1051 TGCAATTAAA CTCGGCCCAA TCTTTTACTA AAAGGATTGA GCCGAATACA

1101 ACAAAGATTC TATTGCATAT ATTTTGACTA AGTATATACT TACCTAGATA
                                         HindIII
1151 TACAAGATTT GAAATACAAA ATCTAGcaag ctt
```

Figure 9

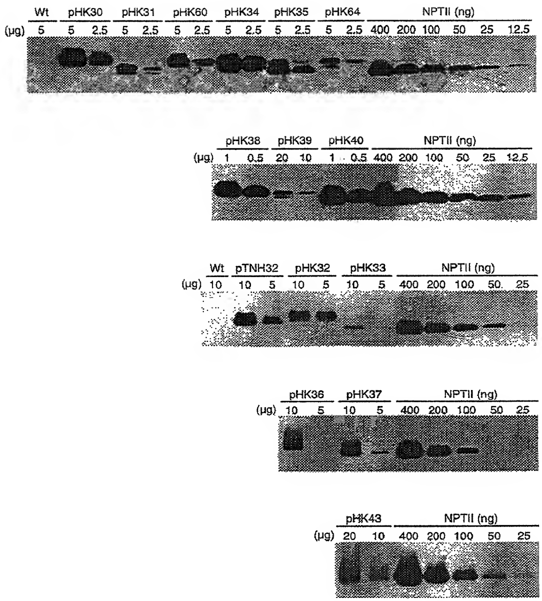

Wt pHK30 pHK31 pHK60 pHK34 pHK35 pHK64 NPTII (ng)
(µg) 5 5 2.5 5 2.5 5 2.5 5 2.5 5 2.5 5 2.5 400 200 100 50 25 12.5
pHK38 pHK39 pHK40 NPTII (ng)
(µg) 1 0.5 20 10 1 0.5 400 200 100 50 25 12.5
Wt pTNH32 pHK32 pHK33 NPTII (ng)
(µg) 10 10 5 10 5 10 5 400 200 100 50 25
pHK36 pHK37 NPTII (ng)
(µg) 10 5 10 5 400 200 100 50 25
pHK43 NPTII (ng)
(µg) 20 10 400 200 100 50 25

Figure 10

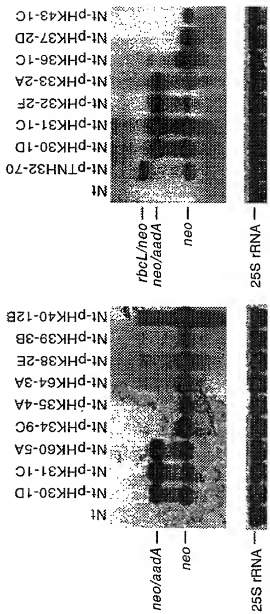
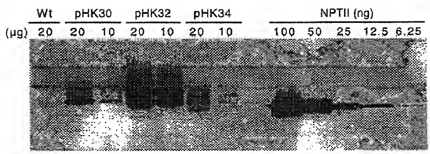

neo/aadA
neo
25S rRNA
Nt
Nt-pHK30-1D
Nt-pHK31-1C
Nt-pHK60-5A
Nt-pHK34-9C
Nt-pHK35-4A
Nt-pHK64-3A
Nt-pHK38-2E
Nt-pHK39-3B
Nt-pHK40-12B
rbcL/neo
neo/aadA
neo
25S rRNA
Nt
Nt-pTNH32-70
Nt-pHK30-1D
Nt-pHK31-1C
Nt-pHK32-2F
Nt-pHK33-2A
Nt-pHK36-1C
Nt-pHK37-2D
Nt-pHK43-1C

Figure 11

atpB wt	**AUG**	AGA	AUC	AAU	CCU	ACU	ACU	UCU	GGU	UCU	GGG	GUU	UCC	ACG
	Met	Arg	Ile	Asn	Pro	Thr	Thr	Ser	Gly	Ser	Gly	Val	Ser	Thr
Fraction	1.0	0.22	0.27	0.61	0.30	0.37	0.37	0.31	0.38	0.31	0.26	0.35	0.14	0.15
Triplet/1000	24.6	7.8	15.5	18.1	13.5	18.4	18.4	20.2	28.2	20.2	19.2	24.9	9.1	7.5
atpB m	**AUG**	AGA	AUa	AAc	CCg	ACa	ACa	agU	Gga	agU	GGG	GUg	UCC	ACG
	Met	Arg	Ile	Asn	Pro	Thr	Thr	Ser	Gly	Ser	Gly	Val	Ser	Thr
Fraction	1.0	0.22	0.29	0.39	0.30	0.23	0.23	0.14	0.24	0.14	0.26	0.21	0.14	0.15
Triplet/1000	24.6	7.8	16.6	11.4	13.2	11.7	11.7	9.3	17.9	9.3	19.2	15.3	9.1	7.5
rbcL wt	**AUG**	UCA	CCA	CAA	ACA	GAG	ACU	AAA	GCA	AGU	GUU	GGA	UUC	AAA
	Met	Ser	Pro	Gln	Thr	Glu	Thr	Lys	Ala	Ser	Val	Gly	Phe	Lys
Fraction	1.0	0.21	0.24	0.57	0.23	0.38	0.37	0.60	0.29	0.14	0.35	0.24	0.40	0.60
Triplet/1000	24.6	13.5	10.6	21.0	11.7	12.4	18.4	22.0	18.1	9.3	24.9	17.9	22.5	22.0
rbcL m	**AUG**	agu	CCu	CAg	ACA	GAa	ACa	AAA	GCc	uca	GUa	GGA	UUC	AAA
	Met	Ser	Pro	Gln	Thr	Glu	Thr	Lys	Ala	Ser	Val	Gly	Phe	Lys
Fraction	1.0	0.14	0.30	0.43	0.23	0.62	0.23	0.60	0.16	0.21	0.31	0.24	0.40	0.60
Triplet/1000	24.6	9.3	13.5	15.5	11.7	20.7	11.7	22.0	10.1	13.5	21.8	17.9	22.5	22.0
T7g10+DB/*Ec*	**AUG**	GCa	AGC	AUG	ACU	GGU	GGA	CAG	gcu	agc	auu	gaa	caa	gau
	Met	Ala	Ser	Met	Thr	Gly	Gly	Gln	Ala	Ser	Ile	Glu	Gln	Asp
Fraction	1.0	0.29	0.07	1.00	0.37	0.38	0.24	0.43	0.39	0.07	0.45	0.62	0.57	0.75
Triplet/1000	24.6	18.1	4.7	24.6	18.4	28.2	17.9	15.5	24.4	4.7	25.9	20.7	21.0	24.6
T7g10+DB/pt	**AUG**	GCa	Auc	acu	agc	ccu	gcc	uuG	gcu	agc	auu	gaa	caa	gau
	Met	Ala	Ile	Thr	Ser	Pro	Ala	Leu	Ala	Ser	Ile	Glu	Gln	Asp
Fraction	1.0	0.29	0.27	0.37	0.07	0.30	0.16	0.24	0.39	0.07	0.45	0.62	0.57	0.75
Triplet/1000	24.6	18.1	15.5	18.4	4.7	13.5	10.1	34.7	24.4	4.7	25.9	20.7	21.0	24.6
T7g10-DB	**AUG**	gcu	agc	auu	gaa	caa	gau	gga	uug	cac	gca	ggu	ucu	ccg
	Met	Ala	Ser	Ile	Glu	Gln	Asp	Gly	Leu	His	Ala	Gly	Ser	Pro
Fraction	1.0	0.39	0.07	0.45	0.62	0.57	0.75	0.24	0.24	0.28	0.29	0.38	0.31	0.30
Triplet/1000	24.6	24.4	4.7	25.9	20.7	21.0	24.6	17.9	34.7	9.1	18.1	28.2	20.2	13.2

Figure 12


Wt
pHK30
pHK32
pHK34
NPTII (ng)
(µg) 20 20 10 20 10 20 10 100 50 25 12.5 6.25

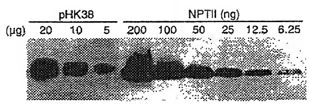

pHK38
NPTII (ng)
(µg) 20 10 5 200 100 50 25 12.5 6.25

Figure 13A

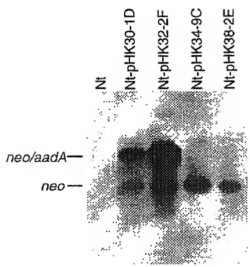


Figure 13B

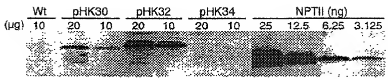

Wt
pHK30
pHK32
pHK34
NPTII (ng)
(μg)
10
20
10
20
10
20
10
25
12.5
6.25
3.125

Figure 14

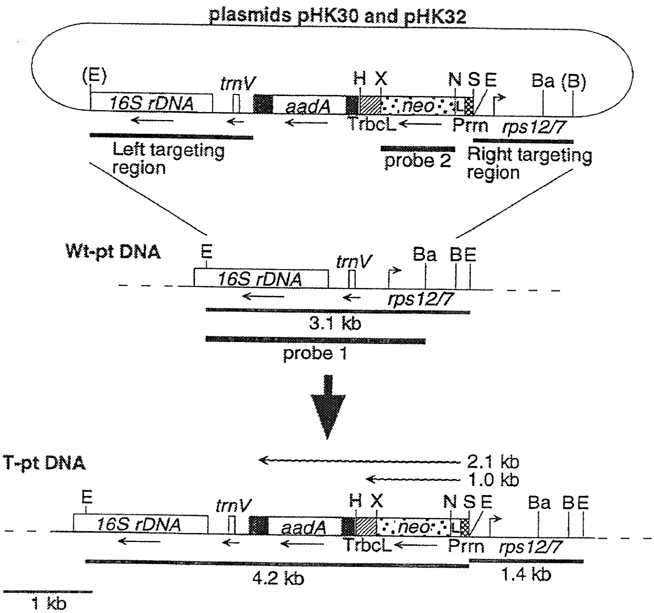

plasmids pHK30 and pHK32
(E)
trnV
H X
N S E
Ba (B)
16S rDNA
aadA
neo
L
TrbcL
Prrn
rps12/7
Left targeting region
probe 2
Right targeting region
Wt-pt DNA
E
trnV
Ba BE
16S rDNA
rps12/7
3.1 kb
probe 1
T-pt DNA
2.1 kb
1.0 kb
E
trnV
H X
N S E
Ba BE
16S rDNA
aadA
neo
L
TrbcL
Prrn
rps12/7
4.2 kb
1.4 kb
1 kb

Figure 15A

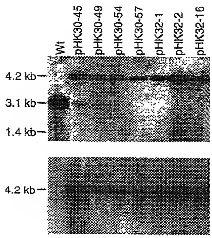

Wt
pHK30-45
pHK30-49
pHK30-54
pHK30-57
pHK32-1
pHK32-2
pHK32-16
4.2 kb
3.1 kb
1.4 kb
4.2 kb

Figure 15B

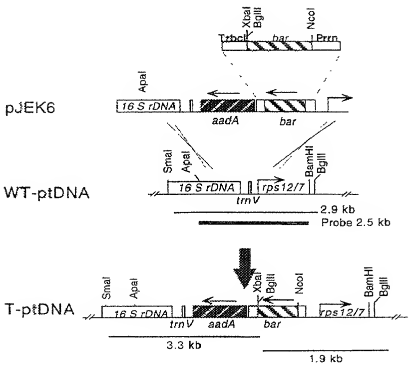

XbaI
BglII
NcoI
Trbcl
bar
Prrn
ApaI
pJEK6
16 S rDNA
aadA
bar
SmaI
ApaI
BamHI
BglII
WT-ptDNA
16 S rDNA
rps12/7
trnV
2.9 kb
Probe 2.5 kb
SmaI
ApaI
XbaI
BglII
NcoI
BamHI
BglII
T-ptDNA
16 S rDNA
trnV
aadA
bar
rps12/7
3.3 kb
1.9 kb

Figure 16A

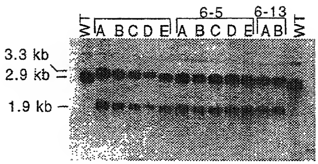


Figure 16B

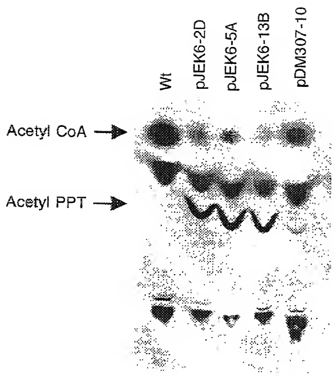

Wt
pJEK6-2D
pJEK6-5A
pJEK6-13B
pDM307-10
Acetyl CoA
Acetyl PPT

Figure 17

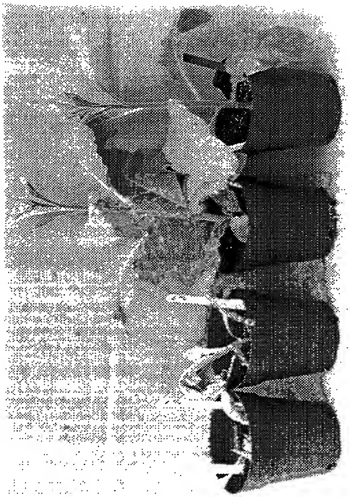


Figure 18A

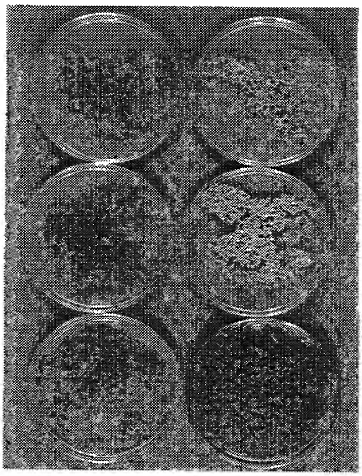


Figure 18B

```
NcoI
CCATGgcaccacaaacagagAGCCCAGAACGACGCCCGGCCGACATCCGCCGTGCCACCG
---------+---------+---------+---------+---------+---------+ 60
GGTACcgtggtgtttgtctcTCGGGTCTTGCTGCGGGCCGGCTGTAGGCGGCACGGTGGC
  M  A  P  Q  T  E  S  P  E  R  R  P  A  D  I  R  R  A  T  E

AGGCGGACATGCCGGCGGTCTGCACCATCGTCAACCACTACATCGAGACAAGCACGGTCA
---------+---------+---------+---------+---------+---------+ 120
TCCGCCTGTACGGCCGCCAGACGTGGTAGCAGTTGGTGATGTAGCTCTGTTCGTGCCAGT
  A  D  M  P  A  V  C  T  I  V  N  H  Y  I  E  T  S  T  V  N

ACTTCCGTACCGAGCCGCAGGAACCGCAGGAGTGGACGGACGACCTCGTCCGTCTGCGGG
---------+---------+---------+---------+---------+---------+ 180
TGAAGGCATGGCTCGGCGTCCTTGGCGTCCTCACCTGCCTGCTGGAGCAGGCAGACGCCC
  F  R  T  E  P  Q  E  P  Q  E  W  T  D  D  L  V  R  L  R  E

AGCGCTATCCCTGGCTCGTCGCCGAGGTGGACGGCGAGGTCGCCGGCATCGCCTACGCGG
---------+---------+---------+---------+---------+---------+ 240
TCGCGATAGGGACCGAGCAGCGGCTCCACCTGCCGCTCCAGCGGCCGTAGCGGATGCGCC
  R  Y  P  W  L  V  A  E  V  D  G  E  V  A  G  I  A  Y  A  G

GCCCCTGGAAGGCACGCAACGCCTACGACTGGACGGCCGAGTCGACCGTGTACGTCTCCC
---------+---------+---------+---------+---------+---------+ 300
CGGGGACCTTCCGTGCGTTGCGGATGCTGACCTGCCGGCTCAGCTGGCACATGCAGAGGG
  P  W  K  A  R  N  A  Y  D  W  T  A  E  S  T  V  Y  V  S  P

CCCGCCACCAGCGGACGGGACTGGGCTCCACGCTCTACACCCACCTGCTGAAGTCCCTGG
---------+---------+---------+---------+---------+---------+ 360
GGGCGGTGGTCGCCTGCCCTGACCCGAGGTGCGAGATGTGGGTGGACGACTTCAGGGACC
  R  H  Q  R  T  G  L  G  S  T  L  Y  T  H  L  L  K  S  L  E

AGGCACAGGGCTTCAAGAGCGTGGTCGCTGTCATCGGGCTGCCCAACGACCCGAGCGTGC
---------+---------+---------+---------+---------+---------+ 420
TCCGTGTCCCGAAGTTCTCGCACCAGCGACAGTAGCCCGACGGGTTGCTGGGCTCGCACG
  A  Q  G  F  K  S  V  V  A  V  I  G  L  P  N  D  P  S  V  R

GCATGCACGAGGCGCTCGGATATGCCCCCCGCGGCATGCTGCGGGCGGCCGGCTTCAAGC
---------+---------+---------+---------+---------+---------+ 480
CGTACGTGCTCCGCGAGCCTATACGGGGGGCGCCGTACGACGCCCGCCGGCCGAAGTTCG
  M  H  E  A  L  G  Y  A  P  R  G  M  L  R  A  A  G  F  K  H

ACGGGAACTGGCATGACGTGGGTTTCTGGCAGCTGGACTTCAGCCTGCCGGTACCGCCCC
---------+---------+---------+---------+---------+---------+ 540
TGCCCTTGACCGTACTGCACCCAAAGACCGTCGACCTGAAGTCGGACGGCCATGGCGGGG
  G  N  W  H  D  V  G  F  W  Q  L  D  F  S  L  P  V  P  P  R

                         BglII
GTCCGGTCCTGCCCGTCACCGAGATCTGATGAtcgaattcctgcagcccgggggatccac
---------+---------+---------+---------+---------+---------+ 600
CAGGCCAGGACGGGCAGTGGCTCTAGACTACTagcttaaggacgtcgggccccctaggtg
  P  V  L  P  V  T  E  I  *

    XbaI
tagttctaga
---------+ 610
atcaagatct
```

Figure 19

```
 NcoI  NheI
CcATGGctAGCCCAGAAaGAaGaCCGGCCGAtATtaGaCGTGCtACaGAaGCtGAtATGC
---------+---------+---------+---------+---------+---------+
ggTACcgaTCGGGTCTTtCTtCtGGCCGGCTaTAatCtGCACGaTGtCTtCGaCTaTACG
 M  A  S  P  E  R  R  P  A  D  I  R  R  A  T  E  A  D  M  P

CaGCaGTtTGtACaATtGTtAAtCAtTAtATaGAaACAAGtACcGTaAACTTtcGaACtG
---------+---------+---------+---------+---------+---------+
GtCGtCAaACaTGtTAaCAaTTaGTaATaTAtCTtTGTTCaTGgCAtTTGAAagCtTGaC
 A  V  C  T  I  V  N  H  Y  I  E  T  S  T  V  N  F  R  T  E

AaCCtCAaGAACCtCAaGAaTGGACtGAtGAttTaGTCCGTtTaCGaGAGCGCTATCCtT
---------+---------+---------+---------+---------+---------+
TtGGaGTtCTTGGaGTtCTtACCTGaCTaCTaaAtCAGGCAaAtGCtCTCGCGATAGGaA
 P  Q  E  P  Q  E  W  T  D  D  L  V  R  L  R  E  R  Y  P  W

GGCTtGTaGCaGAaGTtGACGGaGAaGTaGCtGGgATtGCaTAtGCGGGCCCgTGGAAaG
---------+---------+---------+---------+---------+---------+
CCGAaCAtCGtCTtCAaCTGCCtCTtCAtCGaCCcTAaCGtATaCGCCCGGGcACCTTtC
 L  V  A  E  V  D  G  E  V  A  G  I  A  Y  A  G  P  W  K  A

CAcGaAAtGCaTAtGAtTGGACgGCtGAaTCaACtGTgTACGTtTCaCCaCGtCAtCAaC
---------+---------+---------+---------+---------+---------+
GTgCtTTaCGtATaCTaACCTGcCGaCTtAGtTGaCAcATGCAaAGtGGtGCaGTaGTtG
 R  N  A  Y  D  W  T  A  E  S  T  V  Y  V  S  P  R  H  Q  R

GgACaGGACTtGGtTCtACttTaTAtACcCAtCTaCTGAAaTCttTGGAGGCACAgGGtT
---------+---------+---------+---------+---------+---------+
CcTGtCCTGAaCCaAGaTGaaAtATaTGgGTaGAtGACTTtAGaaACCTCCGTGTcCCaA
 T  G  L  G  S  T  L  Y  T  H  L  L  K  S  L  E  A  Q  G  F

TtAAGAGtGTgGTaGCTGTtATaGGatTGCCgAAtGAtCCctcgGTaCGCATGCAcGAaG
---------+---------+---------+---------+---------+---------+
AaTTCTCaCAcCAtCGACAaTAtCCtaACGGcTTaCTaGGgagcCAtGCGTACGTgCTtC
 K  S  V  V  A  V  I  G  L  P  N  D  P  S  V  R  M  H  E  A

CtCTcGGATATGCtCCcaGaGGtATGtTGaGGGCcGCaGGtTTCAAaCAtGGaAAtTGGC
---------+---------+---------+---------+---------+---------+
GaGAgCCTATACGaGGgtCtCCaTACaACtCCCGgCGtCCaAAGTTtGTaCCtTTaACCG
 L  G  Y  A  P  R  G  M  L  R  A  A  G  F  K  H  G  N  W  H

ATGAtGTaGGTTTtTGGCAaCTtGAcTTCtcttTaCCaGTACCtCCtCGTCCcGTttTaC
---------+---------+---------+---------+---------+---------+
TACTaCAtCCAAAaACCGTtGAaCTgAAGagaaAtGGtCATGGaGGaGCAGGgCAaaAtG
 D  V  G  F  W  Q  L  D  F  S  L  P  V  P  P  R  P  V  L  P

          BglII      XbaI
CcGTtACtGAGATCTGATGAtctaga
---------+---------+------
GgCAaTGaCTCTAGACTACTagatct
 V  T  E  I  *  *
```

Figure 20A

```
NcoI  NheI
ccATGgctAGCCCAGAAaGAaGaCCGGCCGAtATtaGaCGTGCtACaGAaGCtGAtATGC
---------+---------+---------+---------+---------+---------+
ggTACcgaTCGGGTCTTtCTtCtGGCCGGCTaTAatCtGCACGaTGtCTtCGaCTaTACG
  M  A  S  P  E  R  R  P  A  D  I  R  R  A  T  E  A  D  M  P

CaGCaGTtTGtACaATtGTtAAtCAtTAtATaGAaACAAGtACaGTaAAtTTtcGaACtG
---------+---------+---------+---------+---------+---------+
GtCGtCAaACaTGtTAaCAaTTaGTaATaTAtCTtTGTTCaTGtCAtTTaAAagCtTGaC
  A  V  C  T  I  V  N  H  Y  I  E  T  S  T  V  N  F  R  T  E

AaCCtCAaGAACCtCAaGAaTGGACtGAtGAttTaGTaCGTtTaCGaGAaCGtTATCCtT
---------+---------+---------+---------+---------+---------+
TtGGaGTtCTTGGaGTtCTtACCTGaCTaCTaaAtCAtGCAaAtGCtCTtGCaATAGGaA
  P  Q  E  P  Q  E  W  T  D  D  L  V  R  L  R  E  R  Y  P  W

GGCTtGTaGCaGAaGTtGAcGGaGAaGTaGCtGGaATtGCaTAtGCtGGtCCgTGGAAaG
---------+---------+---------+---------+---------+---------+
CCGAaCAtCGtCTtCAaCTgCCtCTtCAtCGaCCtTAaCGtATaCGaCCaGGcACCTTtC
  L  V  A  E  V  D  G  E  V  A  G  I  A  Y  A  G  P  W  K  A

CAcGaAAtGCaTAtGAtTGGACaGCtGAaTCaACtGTtTAtGTtTCaCCaCGtCAtCAaC
---------+---------+---------+---------+---------+---------+
GTgCtTTaCGtATaCTaACCTGtCGaCTtAGtTGaCAaATaCAaAGtGGtGCaGTaGTtG
  R  N  A  Y  D  W  T  A  E  S  T  V  Y  V  S  P  R  H  Q  R

GtACaGGACTtGGtTCtACttTaTAtACtCAtCTtCTtAAaTCttTGGAaGCACAaGGtT
---------+---------+---------+---------+---------+---------+
CaTGtCCTGAaCCaAGaTGaaAtATaTGaGTaGAaGAaTTtAGaaACCTtCGTGTtCCaA
  T  G  L  G  S  T  L  Y  T  H  L  L  K  S  L  E  A  Q  G  F

TtAAaAGtGTaGTaGCTGTtATaGGatTGCCgAAtGAtCCctcaGTaCGCATGCAtGAaG
---------+---------+---------+---------+---------+---------+
AaTTtTCaCAtCAtCGACAaTAtCCtaACGGcTTaCTaGGgagtCAtGCGTACGTaCTtC
  K  S  V  V  A  V  I  G  L  P  N  D  P  S  V  R  M  H  E  A

CtCTtGGATATGCtCCcaGaGGtATGtTGaGGGCaGCaGGtTTCAAaCAtGGaAAtTGGC
---------+---------+---------+---------+---------+---------+
GaGAaCCTATACGaGGgtCtCCaTACaACtCCCGtCGtCCaAAGTTtGTaCCtTTaACCG
  L  G  Y  A  P  R  G  M  L  R  A  A  G  F  K  H  G  N  W  H

ATGAtGTaGGTTTtTGGCAaCTtGAcTTCtcttTaCCaGTACCtCCtCGTCCcGTttTaC
---------+---------+---------+---------+---------+---------+
TACTaCAtCCAAAaACCGTtGAaCTgAAGagaaAtGGtCATGGaGGaGCAGGgCAaaAtG
  D  V  G  F  W  Q  L  D  F  S  L  P  V  P  P  R  P  V  L  P

          BglII       XbaI
CcGTtACtGAGATCTGATGAtctaga
---------+---------+------
GgCAaTGaCTCTAGACTACTagatct
  V  T  E  I  *  *
```

Figure 20B

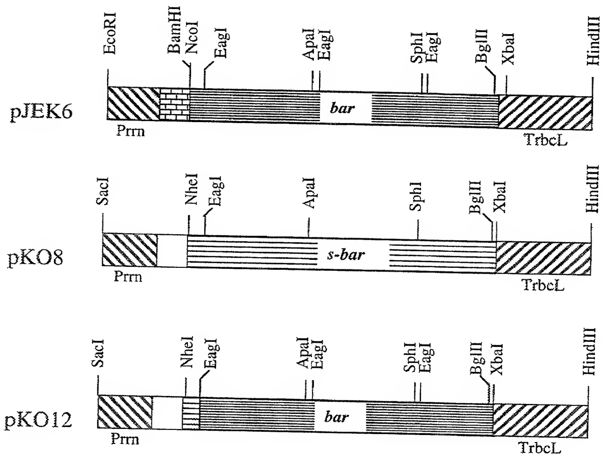

EcoRI
BamHI
NcoI
EagI
ApaI
EagI
SphI
EagI
BglII
XbaI
HindIII
pJEK6
bar
Prrn
TrbcL
SacI
NheI
EagI
ApaI
SphI
BglII
XbaI
HindIII
pKO8
s-bar
Prrn
TrbcL
SacI
NheI
EagI
ApaI
EagI
SphI
EagI
BglII
XbaI
HindIII
pKO12
bar
Prrn
TrbcL

Figure 21

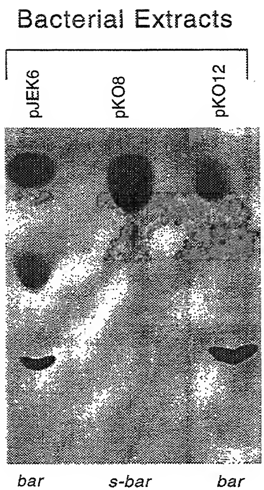


Figure 22A

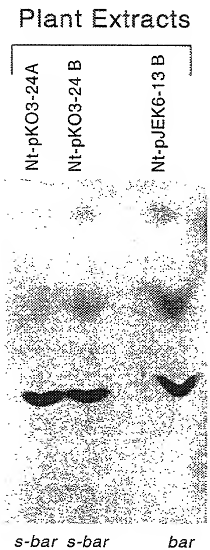

Plant Extracts
Nt-pKO3-24A
Nt-pKO3-24 B
Nt-pJEK6-13 B
s-bar s-bar bar

Figure 22B

Adapter

AAD E L V E G K L E L V E G L K V A GFP

GGTAGTgGGCAAAgaacttgttgaaggaaaattggagctagtagaaggtcttaaagtcgCCATGG

BstXI NcoI

Figure 23A

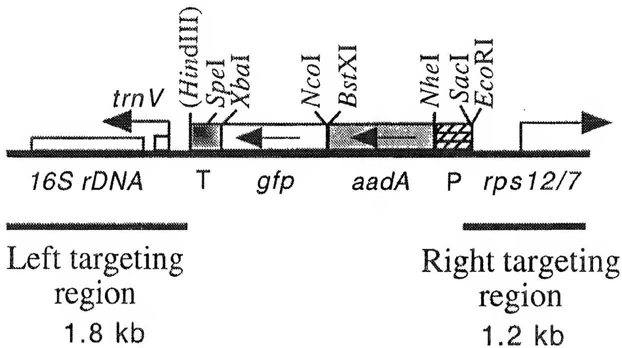


Figure 23B

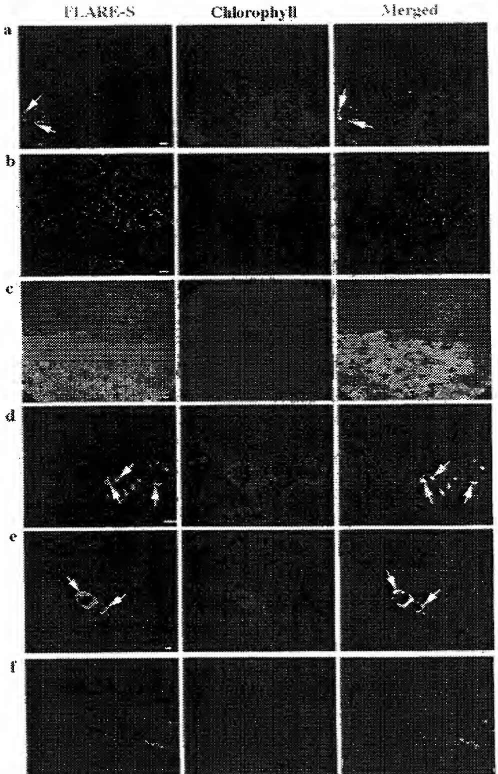


Figure 24

09/762105

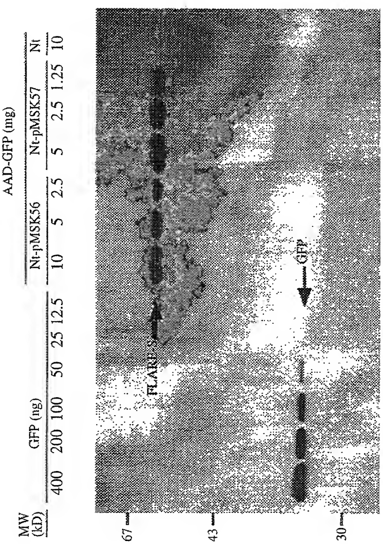


Figure 25

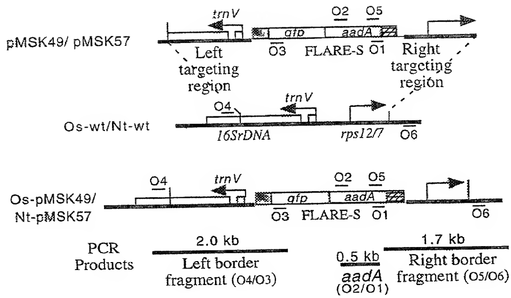


Figure 26A

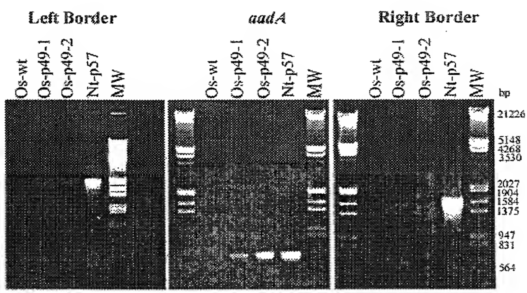


Figure 26B

09/762105

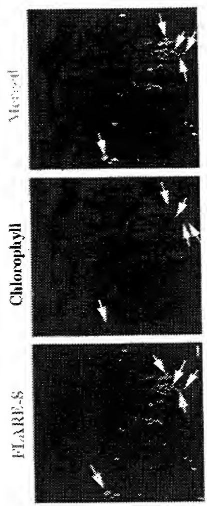


Figure 27

```
FLARE16-S.seq  Length: 1574
     NcoI
   1 cCATGgGGgc tagcGAAGCG GTGATCGCCG AAGTATCGAC TCAACTATCA
  51 GAGGTAGTTG GCGTCATCGA GCGCCATCTC GAACCGACGT TGCTGGCCGT
 101 ACATTTGTAC GGCTCCGCAG TGGATGGCGG CCTGAAGCCA CACAGTGATA
 151 TTGATTTGCT GGTTACGGTG ACCGTAAGGC TTGATGAAAC AACGCGGCGA
 201 GCTTTGATCA ACGACCTTTT GGAAACTTCG GCTTCCCCTG GAGAGAGCGA
 251 GATTCTCCGC GCTGTAGAAG TCACCATTGT TGTGCACGAC GACATCATTC
 301 CGTGGCGTTA TCCAGCTAAG CGCGAACTGC AATTTGGAGA ATGGCAGCGC
 351 AATGACATTC TTGCAGGTAT CTTCGAGCCA GCCACGATCG ACATTGATCT
 401 GGCTATCTTG CTGACAAAAG CAAGAGAACA TAGCGTTGCC TTGGTAGGTC
 451 CAGCGGCGGA GGAACTCTTT GATCCGGTTC CTGAACAGGA TCTATTTGAG  ← aadA
 501 GCGCTAAATG AAACCTTAAC GCTATGGAAC TCGCCGCCCG ACTGGGCTGG
 551 CGATGAGCGA AATGTAGTGC TTACGTTGTC CCGCATTTGG TACAGCGCAG
 601 TAACCGGCAA AATCGCGCCG AAGGATGTCG CTGCCGACTG GGCAATGGAG
 651 CGCCTGCCGG CCCAGTATCA GCCCGTCATA CTTGAAGCTA GACAGGCTTA
 701 TCTTGGACAA GAAGAAGATC GCTTGGCCTC GCGCGCAGAT CAGTTGGAAG
 751 AATTTGTCCA CTACGTGAAA GGCGAGATCA CCAAGGTAGT gGGCAAAgaa
 801 cttgttgaag gaaaattgga gctagtagaa ggtcttaaag tcgccATGgc
 851 tAGTAAAGGA GAAGAACTTT TCACTGGAGT TGTCCCAATT CTTGTTGAAT
 901 TAGATGGTGA TGTTAATGGG CACAAATTTT CTGTCAGTGG AGAGGGTGAA
 951 GGTGATGCAA CATACGGAAA ACTTACCCTT AAATTTATTT GCACTACTGG
1001 AAAACTACCT GTTCCtTGGC CAACACTTGT CACTACTTTC TCTTATGGTG
1051 TTCAATGCTT TTCAAGATAC CCAGATCATA TGAAGCGGCA CGACTTCTTC
1101 AAGAGCGCCA TGCCTGAGGG ATACGTGCAG GAGAGGACCA TCTCTTTCAA
1151 GGACGACGGG AACTACAAGA CACGTGCTGA AGTCAAGTTT GAGGGAGACA
1201 CCCTCGTCAA CAGGATCGAG CTTAAGGGAA TCGATTTCAA GGAGGACGGA  ← gfp
1251 AACATCCTCG GCCACAAGTT GGAATACAAC TACAACTCCC ACAACGTATA
1301 CATCACGGCA GACAAACAAA AGAATGGAAT CAAAGCTAAC TTCAAAATTA
1351 GACACAACAT TGAAGATGGA AGCGTTCAAC TAGCAGACCA TTATCAACAA
1401 AATACTCCAA TTGGCGATGG CCCTGTCCTT TTACCAGACA ACCATTACCT
1451 GTCCACACAA TCTGCCCTTT CGAAAGATCC CAACGAAAAG AGAGACCACA
1501 TGGTCCTTCT TGAGTTTGTA ACAGCTGCTG GGATTACACA TGGCATGGAT
1551 GAACTATACA AATAAGgctc taga
                         XbaI
```

Figure 28

FLARE16-S1.seq Length: 1953

SacI

Latp8DB

1 gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
51 GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG TCGAGTAGAC
101 CTTGTTGTTG TGAaAATTCT TAATTCATGA GTTGTAGGGA GGGATTTATG
151 TCACCACAAA CAGAGACTAA AGCAAGTGTT GGATTCAAAg ctagcGAAGC
201 GGTGATCGCC GAAGTATCGA CTCAACTATC AGAGGTAGTT GGCGTCATCG
251 AGCGCCATCT CGAACCGACG TTGCTGGCCG TACATTTGTA CGGCTCCGCA
301 GTGGATGGCG GCCTGAAGCC ACACAGTGAT ATTGATTTGC TGGTTACGGT
351 GACCGTAAGG CTTGATGAAA CAACGCGGCG AGCTTTGATC AACGACCTTT
401 TGGAAACTTC GGCTTCCCCT GGAGAGAGCG AGATTCTCCG CGCTGTAGAA
451 GTCACCATTG TTGTGCACGA CGACATCATT CCGTGGCGTT ATCCAGCTAA
501 GCGCGAACTG CAATTTGGAG AATGGCAGCG CAATGACATT CTTGCAGGTA
551 TCTTCGAGCC AGCCACGATC GACATTGATC TGGCTATCTT GCTGACAAAA
601 GCAAGAGAAC ATAGCGTTGC CTTGGTAGGT CCAGCGGCGG AGGAACTCTT
651 TGATCCGGTT CCTGAACAGG ATCTATTTGA GGCGCTAAAT GAAACCTTAA
701 CGCTATGGAA CTCGCCGCCC GACTGGGCTG GCGATGAGCG AAATGTAGTG
751 CTTACGTTGT CCCGCATTTG GTACAGCGCA GTAACCGGCA AAATCGCGCC
801 GAAGGATGTC GCTGCCGACT GGGCAATGGA GCGCCTGCCG GCCCAGTATC
851 AGCCCGTCAT ACTTGAAGCT AGACAGGCTT ATCTTGGACA AGAAGAAGAT
901 CGCTTGGCCT CGCGCGCAGA TCAGTTGGAA GAATTTGTCC ACTACGTGAA
951 AGGCGAGATC ACCAAGGTAG TgGGCAAAga acttgttgaa ggaaaattgg
1001 agctagtaga aggtcttaaa gtcgccATGG CTAGTAAAGG AGAAGAACTT
1051 TTCACTGGAG TTGTCCCAAT TCTTGTTGAA TTAGATGGTG ATGTTAATGG
1101 GCACAAATTT TCTGTCAGTG GAGAGGGTGA AGGTGATGCA ACATACGGAA
1151 AACTTACCCT TAAATTTATT TGCACTACTG GAAAACTACC TGTTCCtTGG
1201 CCAACACTTG TCACTACTTT CTCTTATGGT GTTCAATGCT TTTCAAGATA
1251 CCCAGATCAT ATGAAGCGGC ACGACTTCTT CAAGAGCGCC ATGCCTGAGG
1301 GATACGTGCA GGAGAGGACC ATCTCTTTCA AGGACGACGG GAACTACAAG
1351 ACACGTGCTG AAGTCAAGTT TGAGGGAGAC ACCCTCGTCA ACAGGATCGA
1401 GCTTAAGGGA ATCGATTTCA AGGAGGACGG AAACATCCTC GGCCACAAGT
1451 TGGAATACAA CTACAACTCC CACAACGTAT ACATCACGGC AGACAAACAA
1501 AAGAATGGAA TCAAAGCTAA CTTCAAAATT AGACACAACA TTGAAGATGG
1551 AAGCGTTCAA CTAGCAGACC ATTATCAACA AAATACTCCA ATTGGCGATG
1601 GCCCTGTCCT TTTACCAGAC AACCATTACC TGTCCACACA ATCTGCCCTT
1651 TCGAAAGATC CCAACGAAAA GAGAGACCAC ATGGTCCTTC TTGAGTTTGT
1701 AACAGCTGCT GGGATTACAC ATGGCATGGA TGAACTATAC AAATAAGgct
1751 ctagagcGAT CCTGGCCTAG TCTATAGGAG GTTTTGAAAA GAAAGGAGCA
1801 ATAATCATTT TCTTGTTCTA TCAAGAGGGT GCTATTGCTC CTTTCTTTTT
1851 TTCTTTTTAT TTATTTACTA GTATTTTACT TACATAGACT TTTTTGTTTA
1901 CATTATAGAA AAAGAAGGAG AGGTTATTTT CTTGCATTTA TTCATGaaag
1951 ctt

aadA

gfp

TpsbA

HindIII

Figure 29

```
FLARE16-S2.seq  Length: 1985
      SacI
   1  gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG
  51  GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG AATTAACCGA   LrbcL DB
 101  TCGACGTGCa AGCGGACATT TATTTTaAAT TCGATAATTT TTGCAAAAAC
 151  ATTTCGACAT ATTTATTTAT TTTATTATTA TGAGAATCAA TCCTACTACT
 201  TCTGGTTCTG GGGTTTCCAC GgctagcGAA GCGGTGATCG CCGAAGTATC
 251  GACTCAACTA TCAGAGGTAG TTGGCGTCAT CGAGCGCCAT CTCGAACCGA
 301  CGTTGCTGGC CGTACATTTG TACGGCTCCG CAGTGGATGG CGGCCTGAAG
 351  CCACACAGTG ATATTGATTT GCTGGTTACG GTGACCGTAA GGCTTGATGA
 401  AACAACGCGG CGAGCTTTGA TCAACGACCT TTTGGAAACT TCGGCTTCCC
 451  CTGGAGAGAG CGAGATTCTC CGCGCTGTAG AAGTCACCAT TGTTGTGCAC
 501  GACGACATCA TTCCGTGGCG TTATCCAGCT AAGCGCGAAC TGCAATTTGG
 551  AGAATGGCAG CGCAATGACA TTCTTGCAGG TATCTTCGAG CCAGCCACGA
 601  TCGACATTGA TCTGGCTATC TTGCTGACAA AAGCAAGAGA ACATAGCGTT   aadA
 651  GCCTTGGTAG GTCCAGCGGC GGAGGAACTC TTTGATCCGG TTCCTGAACA
 701  GGATCTATTT GAGGCGCTAA ATGAAACCTT AACGCTATGG AACTCGCCGC
 751  CCGACTGGGC TGGCGATGAG CGAAATGTAG TGCTTACGTT GTCCCGCATT
 801  TGGTACAGCG CAGTAACCGG CAAAATCGCG CCGAAGGATG TCGCTGCCGA
 851  CTGGGCAATG GAGCGCCTGC CGGCCCAGTA TCAGCCCGTC ATACTTGAAG
 901  CTAGACAGGC TTATCTTGGA CAAGAAGAAG ATCGCTTGGC CTCGCGCGCA
 951  GATCAGTTGG AAGAATTTGT CCACTACGTG AAAGGCGAGA TCACCAAGGT
1001  AGTgGGCAAA gaacttgttg aaggaaaatt ggagctagta gaaggtctta
1051  aagtcgccAT GgctAGTAAA GGAGAAGAAC TTTTCACTGG AGTTGTCCCA
1101  ATTCTTGTTG AATTAGATGG TGATGTTAAT GGGCACAAAT TTTCTGTCAG
1151  TGGAGAGGGT GAAGGTGATG CAACATACGG AAAACTTACC CTTAAATTTA
1201  TTTGCACTAC TGGAAAACTA CCTGTTCCtT GGCCAACACT TGTCACTACT
1251  TTCTCTTATG GTGTTCAATG CTTTTCAAGA TACCCAGATC ATATGAAGCG
1301  GCACGACTTC TTCAAGAGCG CCATGCCTGA GGGATACGTG CAGGAGAGGA
1351  CCATCTCTTT CAAGGACGAC GGGAACTACA AGACACGTGC TGAAGTCAAG
1401  TTTGAGGGAG ACACCCTCGT CAACAGGATC GAGCTTAAGG GAATCGATTT
1451  CAAGGAGGAC GGAAACATCC TCGGCCACAA GTTGGAATAC AACTACAACT   gfp
1501  CCCACAACGT ATACATCACG GCAGACAAAC AAAAGAATGG AATCAAAGCT
1551  AACTTCAAAA TTAGACACAA CATTGAAGAT GGAAGCGTTC AACTAGCAGA
1601  CCATTATCAA CAAAATACTC CAATTGGCGA TGGCCCTGTC CTTTTACCAG
1651  ACAACCATTA CCTGTCCACA CAATCTGCCC TTTCGAAAGA TCCCAACGAA
1701  AAGAGAGACC ACATGGTCCT TCTTGAGTTT GTAACAGCTG CTGGGATTAC
1751  ACATGGCATG GATGAACTAT ACAAATAAGg ctctagagcG ATCCTGGCCT
1801  AGTCTATAGG AGGTTTTGAA AAGAAAGGAG CAATAATCAT TTTCTTGTTC
1851  TATCAAGAGG GTGCTATTGC TCCTTTCTTT TTTTCTTTTT ATTTATTTAC   TpsbA
1901  TAGTATTTTA CTTACATAGA CTTTTTTGTT TACATTATAG AAAAAGAAGG
1951  AGAGGTTATT TTCTTGCATT TATTCATGaa agctt
                                     HindIII
```

Figure 30

FLARE11-S.seq Length: 1595

```
      NcoI                     C-Myc
   1  ccatggggc  tagcgaacaa aaactcattt ctgaagaaga cttgcctagc
  51  GAAGCGGTGA TCGCCGAAGT ATCGACTCAA CTATCAGAGG TAGTTGGCGT
 101  CATCGAGCGC CATCTCGAAC CGACGTTGCT GGCCGTACAT TTGTACGGCT
 151  CCGCAGTGGA TGGCGGCCTG AAGCCACACA GTGATATTGA TTTGCTGGTT
 201  ACGGTGACCG TAAGGCTTGA TGAAACAACG CGGCGAGCTT TGATCAACGA
 251  CCTTTTGGAA ACTTCGGCTT CCCCTGGAGA GAGCGAGATT CTCCGCGCTG
 301  TAGAAGTCAC CATTGTTGTG CACGACGACA TCATTCCGTG GCGTTATCCA
 351  GCTAAGCGCG AACTGCAATT TGGAGAATGG CAGCGCAATG ACATTCTTGC
 401  AGGTATCTTC GAGCCAGCCA CGATCGACAT TGATCTGGCT ATCTTGCTGA
 451  CAAAAGCAAG AGAACATAGC GTTGCCTTGG TAGGTCCAGC GGCGGAGGAA   aadA
 501  CTCTTTGATC CGGTTCCTGA ACAGGATCTA TTTGAGGCGC TAAATGAAAC
 551  CTTAACGCTA TGGAACTCGC CGCCCGACTG GGCTGGCGAT GAGCGAAATG
 601  TAGTGCTTAC GTTGTCCCGC ATTTGGTACA GCGCAGTAAC CGGCAAAATC
 651  GCGCCGAAGG ATGTCGCTGC CGACTGGGCA ATGGAGCGCC TGCCGGCCCA
 701  GTATCAGCCC GTCATACTTG AAGCTAGACA GGCTTATCTT GGACAAGAAG
 751  AAGATCGCTT GGCCTCGCGC GCAGATCAGT TGGAAGAATT TGTCCACTAC
 801  GTGAAAGGCG AGATCACCAA GGTAGTgGGC AAAgaacttg cagttgaagg
 851  aaaattggag gtcgccATGg ctAGTAAAGG AGAAGAACTT TTCACTGGAG
 901  TTGTCCCAAT TCTTGTTGAA TTAGATGGTG ATGTTAATGG GCACAAATTT
 951  TCTGTCAGTG GAGAGGGTGA AGGTGATGCA ACATACGGAA AACTTACCCT
1001  TAAATTTATT TGCACTACTG GAAAACTACC TGTTCCtTGG CCAACACTTG
1051  TCACTACTTT CTCTTATGGT GTTCAATGCT TTTCAAGATA CCCAGATCAT
1101  ATGAAGCGGC ACGACTTCTT CAAGAGCGCC ATGCCTGAGG GATACGTGCA
1151  GGAGAGGACC ATCTCTTTCA AGGACGACGG GAACTACAAG ACACGTGCTG
1201  AAGTCAAGTT TGAGGGAGAC ACCCTCGTCA ACAGGATCGA GCTTAAGGGA   gfp
1251  ATCGATTTCA AGGAGGACGG AAACATCCTC GGCCACAAGT TGGAATACAA
1301  CTACAACTCC CACAACGTAT ACATCACGGC AGACAAACAA AAGAATGGAA
1351  TCAAAGCTAA CTTCAAAATT AGACACAACA TTGAAGATGG AAGCGTTCAA
1401  CTAGCAGACC ATTATCAACA AAATACTCCA ATTGGCGATG GCCCTGTCCT
1451  TTTACCAGAC AACCATTACC TGTCCACACA ATCTGCCCTT TCGAAAGATC
1501  CCAACGAAAA GAGAGACCAC ATGGTCCTTC TTGAGTTTGT AACAGCTGCT
1551  GGGATTACAC ATGGCATGGA TGAACTATAC AAATAAGgct ctaga
                                               XbaI
```

Figure 31

```
FLARE11-S3.seq  Length: 1961

     SacI
   1 gagctcGCTC CCCCGCCGTC GTTCAATGAG AATGGATAAG AGGCTCGTGG   L T7g10DB
  51 GATTGACGTG AGGGGGCAGG GATGGCTATA TTTCTGGGAG GGAGACCACA
 101 ACGGTTTCCC aCTAGAAATA ATTTTGTTTA ACTTTAAGAA GGAGATATAC
 151 ATATGGCaAG CATGACTGGT GGACAGgcta gcgaacaaaa actcatttct
 201 gaagaagact tgcctagcGA AGCGGTGATC GCCGAAGTAT CGACTCAACT
 251 ATCAGAGGTA GTTGGCGTCA TCGAGCGCCA TCTCGAACCG ACGTTGCTGG
 301 CCGTACATTT GTACGGCTCC GCAGTGGATG GCGGCCTGAA GCCACACAGT
 351 GATATTGATT TGCTGGTTAC GGTGACCGTA AGGCTTGATG AAACAACGCG
 401 GCGAGCTTTG ATCAACGACC TTTTGGAAAC TTCGGCTTCC CCTGGAGAGA
 451 GCGAGATTCT CCGCGCTGTA GAAGTCACCA TTGTTGTGCA CGACGACATC
 501 ATTCCGTGGC GTTATCCAGC TAAGCGCGAA CTGCAATTTG GAGAATGGCA
 551 GCGCAATGAC ATTCTTGCAG GTATCTTCGA GCCAGCCACG ATCGACATTG
 601 ATCTGGCTAT CTTGCTGACA AAAGCAAGAG AACATAGCGT TGCCTTGGTA   aadA
 651 GGTCCAGCGG CGGAGGAACT CTTTGATCCG GTTCCTGAAC AGGATCTATT
 701 TGAGGCGCTA AATGAAACCT TAACGCTATG GAACTCGCCG CCCGACTGGG
 751 CTGGCGATGA GCGAAATGTA GTGCTTACGT TGTCCCGCAT TTGGTACAGC
 801 GCAGTAACCG GCAAAATCGC GCCGAAGGAT GTCGCTGCCG ACTGGGCAAT
 851 GGAGCGCCTG CCGGCCCAGT ATCAGCGCGT CATACTTGAA GCTAGACAGG
 901 CTTATCTTGG ACAAGAAGAA GATCGCTTGG CCTCGCGCGC AGATCAGTTG
 951 GAAGAATTTG TCCACTACGT GAAAGGCGAG ATCACCAAGG TAGTgGGCAA
1001 Agaacttgca gttgaaggaa aattggaggt cgccATGgct AGTAAAGGAG
1051 AAGAACTTTT CACTGGAGTT GTCCCAATTC TTGTTGAATT AGATGGTGAT
1101 GTTAATGGGC ACAAATTTTC TGTCAGTGGA GAGGGTGAAG GTGATGCAAC
1151 ATACGGAAAA CTTACCCTTA AATTTATTTG CACTACTGGA AAACTACCTG
1201 TTCCtTGGCC AACACTTGTC ACTACTTTCT CTTATGGTGT TCAATGCTTT
1251 TCAAGATACC CAGATCATAT GAAGCGGCAC GACTTCTTCA AGAGCGCCAT
1301 GCCTGAGGGA TACGTGCAGG AGAGGACCAT CTCTTTCAAG GACGACGGGA
1351 ACTACAAGAC ACGTGCTGAA GTCAAGTTTG AGGGAGACAC CCTCGTCAAC
1401 AGGATCGAGC TTAAGGGAAT CGATTTCAAG GAGGACGGAA ACATCCTCGG   gfp
1451 CCACAAGTTG GAATACAACT ACAACTCCCA CAACGTATAC ATCACGGCAG
1501 ACAAACAAAA GAATGGAATC AAAGCTAACT TCAAAATTAG ACACAACATT
1551 GAAGATGGAA GCGTTCAACT AGCAGACCAT TATCAACAAA ATACTCCAAT
1601 TGGCGATGGC CCTGTCCTTT TACCAGACAA CCATTACCTG TCCACACAAT
1651 CTGCCCTTTC GAAAGATCCC AACGAAAAGA GAGACCACAT GGTCCTTCTT
1701 GAGTTTGTAA CAGCTGCTGG GATTACACAT GGCATGGATG AACTATACAA
1751 ATAAGgctct agagcGATCC TGGCCTAGTC TATAGGAGGT TTTGAAAAGA
1801 AAGGAGCAAT AATCATTTTC TTGTTCTATC AAGAGGGTGC TATTGCTCCT
1851 TTCTTTTTTT CTTTTTATTT ATTTACTAGT ATTTTACTTA CATAGACTTT   TpsbA
1901 TTTGTTTACA TTATAGAAAA AGAAGGAGAG GTTATTTTCT TGCATTTATT
1951 CATGaaagct t
        HindIII
```

Figure 32

pMSK35.seq Length: 4671

Figure 33A

0122488

```
1     GGGAACGGAT TCACCGCCGT ATGGCTGACC GGCGATTACT AGCGATTCCT
51    GCTTCATGCA GGCGAGTTGC AGCCTGCAAT CCGAACTGAG GACGGGTTTT
101   TGGAGTTAGC TCACCCTCGC GAGATCGCGA CCCTTTGTCC CGCCCATTGT
151   AGCACGTGTG TCGCCCAGGG CATAAGGGGC ATGATGACTT GGCCTCATCC
201   TCTCCTTCCT CCGGCTTAAC ACCGGCGGTC TGTTCAGGGT TCCAAACTCA
251   TAGTGGCAAC TAAACACGAG GGTTGCGCTC GTTGCGAGAC TTAACCCAAC
301   ACCTTACGGC ACGAGCTGAC GACAGCCATG CACCACCTGT GTCCGCGTTC
351   CCGAGGGCAC CCCTCTCTTT CAAGAGGATT CGCGGCATGT CAAGCCCTGG
401   TAAGGTTCTT CGCTTTGCAT CGAATTAAAC CACATGCTCC ACCGCTTGTG
451   CGGGCCCCCG TCAATTCCTT TGAGTTTCAT TCTTGCGAAC GTACTCCCCA
501   GGCGGGATAC TTAACGCGTT AGCTACAGCA CTGCACGGGT CGAGTCGCAC
551   AGCACCTAGT ATCCATCGTT TACGGCTAGG ACTACTGGGG TCTCTAATCC
601   CATTTGCTCC CCTAGCTTTC GTCTCTCAGT GTCAGTGTCG GCCCAGCAGA
651   GTGCTTTCGC CGTTGGTGTT CTTTCCGATC TCAATGCATT TCACCGCTCC
701   ACCGGAAATT CCCTCTGCCC CTACCGTACT CCAGCTTGGT AGTTTCCACC
751   GCCTGTCCAG GGTTGAGCCC TGGGATTTGA CGGCGGACTT GAAAAGCCAC
801   CTACAGACGC TTTACGCCCA ATCATTCCGG ATAACGCTTG CATCCTCTGT
851   CTTACCGCGG CTGCTGGCAC AGAGTTAGCC GATGCTTATT CCTCAGATAC
901   CGTCATTGTT TCTTCTCCGA GAAAAGAAGT TGACGACCCG TGGGCCTTCC
951   ACCTCCACGC GGCATTGCTC CGTCAGGCTT TCGCCCATTG CGGAAAATTC
1001  CCCACTGCTG CCTCCCGTAG GAGTCTGGGC CGTGTCTCAG TCCCAGTGTG
1051  GCTGATCATC CTCTCGGACC AGCTACTGAT CATCGCCTTG GTAAGCTATT
1101  GCCTCACCAA CTAGCTAATC AGACGCGAGC CCCTCCTTGG GCGGATTTCT
1151  CCTTTTGCTC CTCAGCCTAC GGGGTATTAG CAACCGTTTC CAGTTGTTGT
1201  TCCCCTCCCA AGGGCAGGTT CTTACGCGTT ACTCACCCGT TCGCCACTGG
1251  AAACACCACT TCCCGTTCGA CTTGCATGTG TTAAGCATGC CGCCAGCGTT
1301  CATCCTGAGC CAGGATCGAA CTCTCCATGA GATTCATAGT TGCATTACTT
1351  ATAGCTTCCT TATTCGTAGA CAAAGCGGAT TCGGAATTGT CTTTCCTTCC
1401  AAGGATAACT TGTATCCATG CGCTTCAGAT TATTAGCCTG GAGTTCGCCA
1451  CCAGCAGTAT AGCCAACCCT ACCCTATCAC GTCAATCCCA CAAGCCTCTT
1501  ATCCATTCCC GTTCGATCGT GGCGGGGGGA GTAAGTCAAA ATAGAAAAAA
1551  CTCACATTGG GTTTAGGGAT AATCAGGCTC GAACTGATGA CTTCCACCAC
1601  GTCAAGGTGA CACTCTACCG CTGAGTTATA TCCCTTCCCC GTCCCCTCGA
1651  GAAAGAGAAT TACCGAATCC TAAGGCAAAG GGGCGAGAAA CTCAAGGCCA
1701  CCCTTCCTCC GGGCTTTCTT TCCACACTAT TATGGATAGT CAAATAATGG
1751  GAAAAATTGG ATTCAATTGT CAACCGGTCC TATCGAAAAT AGGATTGACT
1801  ATGGATTCGA GCCATAGCAC ATGGTTTCAT AAAATCTGTA CGATTTTCCC
1851  GATCTAAATC GAGCAGGTTT CCATGAAGAA gatcgacggt atcgataagc
1901  ttgcatgcct gcaggtCGAA TATAGCTCTT CTTTCTTATT TCAATGATAT
1951  TATTATTTCA AAGATAAGAG ATATTCAAAG ATAAGAGATA AGAAGAAGTC
2001  AAAATTTGAT TTTTTTTTTG GAAAAAAAAA ATCAAAAAGA TATAGTAACA
2051  TTAGCAAGAA GAGAAACAAG TTCTATTTCA CAATTTAAAC AAATACAAAA
2101  TCAAAATAGA ATACTCAATC ATGAATAAAT GCAAGAAAAT AACCTCTCCT
2151  TCTTTTTCTA TAATGTAAAC AAAAAAGTCT ATGTAAGTAA AATACTAGTA
2201  AATAAATAAA AAGAAAAAAA GAAAGGAGCA ATAGCACCCT CTTGATAGAA
2251  CAAGAAAATG ATTATTGCTC CTTTCTTTTC AAAACCTCCT ATAGACTAGG
2301  CCAGGATCgc tctagcTAGA CATTATTTGC CGACTACCTT GGTGATCTCG
2351  CCTTTCACGT AGTGGACAAA TTCTTCCAAC TGATCTGCGC GCGAGGCCAA
2401  GCGATCTTCT TCTTGTCCAA GATAAGCCTG TCTAGCTTCA AGTATGACGG
2451  GCTGATACTG GGCCGGCAGG CGCTCCATTG CCCAGTCGGC AGCGACATCC
2501  TTCGGCGCGA TTTTGCCGGT TACTGCGCTG TACCAAATGC GGGACAACGT
2551  AAGCACTACA TTTCGCTCAT CGCCAGCCCA GTCGGGCGGC GAGTTCCATA
2601  GCGTTAAGGT TTCATTTAGC GCCTCAAATA GATCCTGTTC AGGAACCGGA
2651  TCAAAGAGTT CCTCCGCCGC TGGACCTACC AAGGCAACGC TATGTTCTCT
2701  TGCTTTTGTC AGCAAGATAG CCAGATCAAT GTCGATCGTG GCTGGCTCGA
```

Rice left targeting sequence

aadA

pMSK35.seq Length: 4671

```
2751  AGATACCTGC AAGAATGTCA TTGCGCTGCC ATTCTCCAAA TTGCAGTTCG
2801  CGCTTAGCTG GATAACGCCA CGGAATGATG TCGTCGTGCA CAACAATGGT
2851  GACTTCTACA GCGCGGAGAA TCTCGCTCTC TCCAGGGGAA GCCGAAGTTT
2901  CCAAAAGGTC GTTGATCAAA GCTCGCCGCG TTGTTTCATC AAGCCTTACG
2951  GTCACCGTAA CCAGCAAATC AATATCACTG TGTGGCTTCA GGCCGCCATC
3001  CACTGCGGAG CCGTACAAAT GTACGGCCAG CAACGTCGGT TCGAGATGGC
3051  GCTCGATGAC GCCAACTACC TCTGATAGTT GAGTCGATAC TTCGGCGATC
3101  ACCGCTTCCC TCATGgATCC CTCCCTACAA CTGTATCCAa GCGCTTCgTA
3151  TTCGCCCGGA GTTCGCTCCC AGAAATATAG CCATCCCTGC CCCCTCACGT
3201  CAATCCCACG AGCCTCTTAT CCATTCTCAT TGAACGACGG CGGGGGAGCT
3251  ttgggtaccg agctcgaatt cctgcagccc gatcTTACCA TTTCCGAAGG
3301  AACTGGGGCT ACATTTCTTT TCAATTTCCA TTCAAGAGTT TCTTATCTGT
3351  TTCCACGCCC TTTTTTGAGA CCTCGAAACA TGAAATGGAC AAATTCCTTC
3401  TCTTAGGAAC ACATACAAGA AAAAGGATAA TGGTAGCCCT CCCATTAACT
3451  ACTTCATTTC ATTTATGAAT TTCATAGTAA TAGAAATCCA TGTCCTACCG
3501  AGACAGAATT TCGAACTTGC TATCCTCTTG CCTAATAGGC AAAGATTGAC
3551  CTCTGTAGAA AGAATGATTC ATTCGGATCG ATATGAGGAC CCAACTACGT
3601  TGCATTGCAG AATCCATGTT CCATATTTGA AGAGGGTTGA CCTCTGTGCT
3651  TCTCTCATGG TACAATCCTC TTCCTGCTGA GCCCCCTTTC TCCTCGGTCC
3701  ACAGAGAAAA AATGGAGGAC TGGTGCGGAC AGTTCATCAC GGAAGAAAGA
3751  ACTCACAGAG CCGGGATCGC TAACTAATAG AATAGTACTA CTAACTAATA
3801  CTAATATATA GAAATAGATA TctagctagA AATAGAAACA ACTAATATAT
3851  AGATAATCGA AATTGAAAAG AACTGTCTTT TCTGTATACT TTCCCCGTTC
3901  TATTGCTACC GCGGGTCTTA TGCAATCGAT CGGATCATAT AGATATCCCT
3951  TCAACACAAC ATAGGTCATC GAAAGGATCT CGGACGACTC ACCAAAGCAC
4001  GAAAGCCAGT TAGAAAATGG ATTCCTATTT GAAGAGTGCC TAACCGCATG
4051  GATAAGCTCA CATTAACCCG TCAATTTTGG ATCCAATTCG GGATTTTTCT
4101  TGGGAAGTTT CGGGAAGAAA TTGGAATGGA ATAATATAGA TTCATACAGA
4151  GGAAAAGGTT CTCTATTGAT GCAAACGCTG TACCTAGAGG ATAGGGATAG
4201  AGGAAGAGGG AAAAATCGAA ATGAAATAAA TAAAGAATAA AGCAAAAAAA
4251  AAATAAGTCG AAGATAGAAG AGCCCAGATT CCAAATGAAG AAATGGAAAC
4301  TCGAAAAGGA TCCTTCTGAT TCTCAAAGAA TGAGGGGCAA GGGGATTGAT
4351  ACCGAGAAAG ATTTCTTCTT ATTATAAGAC GTGATTTGAT CCGCATATGT
4401  TTGGTAAAAG AACAATCTTC TCCTTTAATC ATAAATGGAA AGTGTTCAAT
4451  TAGAACATGA AAACGTGACT CAATTGGTCT TAGTTAGTCT TCGGGACGGA
4501  GTGGAAGAAA GGGCGAAGAC TCTCGAACGA GGAAAAGGAT CCCTTCGAAA
4551  GAATTGAACG AGGAGCCGTA TTAGGTGAAA ATCTCATGTA CGATTCTGTA
4601  GAGGGACAGG AAGGGTGACT TATCTGTCGA CTTTTCCACT ATCAACCCCA
4651  AAAAACCCAA CTCTGCCTTA C
```

aadA

Rice Right targeting sequence

125878

Figure 33B

09/762105

pMSK49.seq Length: 5263

Figure 34A

122488

1	GGGAACGGAT	TCACCGCCGT	ATGGCTGACC	GGCGATTACT	AGCGATTCCT
51	GCTTCATGCA	GGCGAGTTGC	AGCCTGCAAT	CCGAACTGAG	GACGGGTTTT
101	TGGAGTTAGC	TCACCCTCGC	GAGATCGCGA	CCCTTTGTCC	CGCCCATTGT
151	AGCACGTGTG	TCGCCCAGGG	CATAAGGGGC	ATGATGACTT	GGCCTCATCC
201	TCTCCTTCCT	CCGGCTTAAC	ACCGGCGGTC	TGTTCAGGGT	TCCAAACTCA
251	TAGTGGCAAC	TAAACACGAG	GGTTGCGCTC	GTTGCGAGAC	TTAACCCAAC
301	ACCTTACGGC	ACGAGCTGAC	GACAGCCATG	CACCACCTGT	GTCCGCGTTC
351	CCGAGGGCAC	CCCTCTCTTT	CAAGAGGATT	CGCGGCATGT	CAAGCCCTGG
401	TAAGGTTCTT	CGCTTTGCAT	CGAATTAAAC	CACATGCTCC	ACCGCTTGTG
451	CGGGCCCCCG	TCAATTCCTT	TGAGTTTCAT	TCTTGCGAAC	GTACTCCCCA
501	GGCGGGATAC	TTAACGCGTT	AGCTACAGCA	CTGCACGGGT	CGAGTCGCAC
551	AGCACCTAGT	ATCCATCGTT	TACGGCTAGG	ACTACTGGGG	TCTCTAATCC
601	CATTTGCTCC	CCTAGCTTTC	GTCTCTCAGT	GTCAGTGTCG	GCCCAGCAGA
651	GTGCTTTCGC	CGTTGGTGTT	CTTTCCGATC	TCAATGCATT	TCACCGCTCC
701	ACCGGAAATT	CCCTCTGCCC	CTACCGTACT	CCAGCTTGGT	AGTTTCCACC
751	GCCTGTCCAG	GGTTGAGCCC	TGGGATTTGA	CGGCGGACTT	GAAAAGCCAC
801	CTACAGACGC	TTTACGCCCA	ATCATTCCGG	ATAACGCTTG	CATCCTCTGT
851	CTTACCGCGG	CTGCTGGCAC	AGAGTTAGCC	GATGCTTATT	CCTCAGATAC
901	CGTCATTGTT	TCTTCTCCGA	GAAAAGAAGT	TGACGACCCG	TGGGCCTTCC
951	ACCTCCACGC	GGCATTGCTC	CGTCAGGCTT	TCGCCCATTG	CGGAAAATTC
1001	CCCACTGCTG	CCTCCCGTAG	GAGTCTGGGC	CGTGTCTCAG	TCCCAGTGTG
1051	GCTGATCATC	CTCTCGGACC	AGCTACTGAT	CATCGCCTTG	GTAAGCTATT
1101	GCCTCACCAA	CTAGCTAATC	AGACGCGAGC	CCCTCCTTGG	GCGGATTTCT
1151	CCTTTTGCTC	CTCAGCCTAC	GGGGTATTAG	CAACCGTTTC	CAGTTGTTGT
1201	TCCCCTCCCA	AGGGCAGGTT	CTTACGCGTT	ACTCACCCGT	TCGCCACTGG
1251	AAACACCACT	TCCCGTTCGA	CTTGCATGTG	TTAAGCATGC	CGCCAGCGTT
1301	CATCCTGAGC	CAGGATCGAA	CTCTCCATGA	GATTCATAGT	TGCATTACTT
1351	ATAGCTTCCT	TATTCGTAGA	CAAAGCGGAT	TCGGAATTGT	CTTTCCTTCC
1401	AAGGATAACT	TGTATCCATG	CGCTTCAGAT	TATTAGCCTG	GAGTTCGCCA
1451	CCAGCAGTAT	AGCCAACCCT	ACCCTATCAC	GTCAATCCCA	CAAGCCTCTT
1501	ATCCATTCCC	GTTCGATCGT	GGCGGGGGGA	GTAAGTCAAA	ATAGAAAAAA
1551	CTCACATTGG	GTTTAGGGAT	AATCAGGCTC	GAACTGATGA	CTTCCACCAC
1601	GTCAAGGTGA	CACTCTACCG	CTGAGTTATA	TCCCTTCCCC	GTCCCCTCGA
1651	GAAAGAGAAT	TACCGAATCC	TAAGGCAAAG	GGGCGAGAAA	CTCAAGGCCA
1701	CCCTTCCTCC	GGGCTTTCTT	TCCACACTAT	TATGGATAGT	CAAATAATGG
1751	GAAAAATTGG	ATTCAATTGT	CAACCGGTCC	TATCGAAAAT	AGGATTGACT
1801	ATGGATTCGA	GCCATAGCAC	ATGGTTTCAT	AAAATCTGTA	CGATTTTCCC
1851	GATCTAAATC	GAGCAGGTTT	CCATGAAGAA	gatcgacggt	atcgataagc
1901	tttCATGAAT	AAATGCAAGA	AAATAACCTC	TCCTTCTTTT	TCTATAATGT
1951	AAACAAAAAA	GTCTATGTAA	GTAAAATACT	AGTAAATAAA	TAAAAAGAAA
2001	AAAAGAAAGG	AGCAATAGCA	CCCTCTTGAT	AGAACAAGAA	AATGATTATT
2051	GCTCCTTTCT	TTTCAAAACC	TCCTATAGAC	TAGGCCAGGA	TCgctctaga
2101	gcCTTATTTG	TATAGTTCAT	CCATGCCATG	TGTAATCCCA	GCAGCTGTTA
2151	CAAACTCAAG	AAGGACCATG	TGGTCTCTCT	TTTCGTTGGG	ATCTTTCGAA
2201	AGGGCAGATT	GTGTGGACAG	GTAATGGTTG	TCTGGTAAAA	GGACAGGGCC
2251	ATCGCCAATT	GGAGTATTTT	GTTGATAATG	GTCTGCTAGT	TGAACGCTTC
2301	CATCTTCAAT	GTTGTGTCTA	ATTTTGAAGT	TAGCTTTGAT	TCCATTCTTT
2351	TGTTTGTCTG	CCGTGATGTA	TACGTTGTGG	GAGTTGTAGT	TGTATTCCAA
2401	CTTGTGGCCG	AGGATGTTTC	CGTCCTCCTT	GAAATCGATT	CCCTTAAGCT
2451	CGATCCTGTT	GACGAGGGTG	TCTCCCTCAA	ACTTGACTTC	AGCACGTGTC
2501	TTGTAGTTCC	CGTCGTCCTT	GAAAGAGATG	GTCCTCTCCT	GCACGTATCC
2551	CTCAGGCATG	GCGCTCTTGA	AGAAGTCGTG	CCGCTTCATA	TGATCTGGGT
2601	ATCTTGAAAA	GCATTGAACA	CCATAAGAGA	AAGTAGTGAC	AAGTGTTGGC
2651	CAaGGAACAG	GTAGTTTTCC	AGTAGTGCAA	ATAAATTTAA	GGGTAAGTTT
2701	TCCGTATGTT	GCATCACCTT	CACCCTCTCC	ACTGACAGAA	AATTTGTGCC

Rice left targeting sequence

TpsbA

gfp

Figure 34B

pMSK49.seq Length: 5263

```
2751 CATTAACATC ACCATCTAAT TCAACAAGAA TTGGGACAAC TCCAGTGAAA
2801 AGTTCTTCTC CTTTACTagc CATggcgacc tccaattttc cttcaactgc
2851 aagttcgTTG CCCACTACCT TGGTGATCTC GCCTTTCACG TAGTGGACAA
2901 ATTCTTCCAA CTGATCTGCG CGCGAGGCCA AGCGATCTTC TTCTTGTCCA
2951 AGATAAGCCT GTCTAGCTTC AAGTATGACG GGCTGATACT GGGCCGGCAG
3001 GCGCTCCATT GCCCAGTCGG CAGCGACATC CTTCGGCGCG ATTTTGCCGG
3051 TTACTGCGCT GTACCAAATG CGGGACAACG TAAGCACTAC ATTTCGCTCA
3101 TCGCCAGCCC AGTCGGGCGG CGAGTTCCAT AGCGTTAAGG TTTCATTTAG
3151 CGCCTCAAAT AGATCCTGTT CAGGAACCGG ATCAAAGAGT TCCTCCGCCG
3201 CTGGACCTAC CAAGGCAACG CTATGTTCTC TTGCTTTTGT CAGCAAGATA
3251 GCCAGATCAA TGTCGATCGT GGCTGGCTCG AAGATACCTG CAAGAATGTC
3301 ATTGCGCTGC CATTCTCCAA ATTGCAGTTC GCGCTTAGCT GGATAACGCC
3351 ACGGAATGAT GTCGTCGTGC ACAACAATGG TGACTTCTAC AGCGCGGAGA
3401 ATCTCGCTCT CTCCAGGGGA AGCCGAAGTT TCCAAAAGGT CGTTGATCAA
3451 AGCTCGCCGC GTTGTTTCAT CAAGCCTTAC GGTCACCGTA ACCAGCAAAT
3501 CAATATCACT GTGTGGCTTC AGGCCGCCAT CCACTGCGGA GCCGTACAAA
3551 TGTACGGCCA GCAACGTCGG TTCGAGATGG CGCTCGATGA CGCCAACTAC
3601 CTCTGATAGT TGAGTCGATA CTTCGGCGAT CACCGCTTCg ctaggcaagt
3651 cttcttcaga aatgagtttt tgttcgctag cCTGTCCACC AGTCATGCTt
3701 GCCATATGTA TATCTCCTTC TTAAAGTTAA ACAAAATTAT TTCTAGtGGG
3751 AAACCGTTGT GGTCTCCCTC CCAGAAATAT AGCCATCCCT GCCCCCTCAC
3801 GTCAATCCCA CGAGCCTCTT ATCCATTCTC ATTGAACGAC GGCGGGGGAG
3851 Cgagctcgaa ttcctgcagc ccgatcTTAC CATTTCCGAA GGAACTGGGG
3901 CTACATTTCT TTTCAATTTC CATTCAAGAG TTTCTTATCT GTTTCCACGC
3951 CCTTTTTTGA GACCTCGAAA CATGAAATGG ACAAATTCCT TCTCTTAGGA
4001 ACACATACAA GAAAAAGGAT AATGGTAGCC CTCCCATTAA CTACTTCATT
4051 TCATTTATGA ATTTCATAGT AATAGAAATC CATGTCCTAC CGAGACAGAA
4101 TTTCGAACTT GCTATCCTCT TGCCTAATAG GCAAAGATTG ACCTCTGTAG
4151 AAAGAATGAT TCATTCGGAT CGATATGAGG ACCCAACTAC GTTGCATTGC
4201 AGAATCCATG TTCCATATTT GAAGAGGTTT GACCTCTGTG CTTCTCTCAT
4251 GGTACAATCC TCTTCCTGCT GAGCCCCCTT TCTCCTCGGT CCACAGAGAA
4301 AAAATGGAGG ACTGGTGCCG ACAGTTCATC ACGGAAGAAA GAACTCACAG
4351 AGCCGGGATC GCTAACTAAT AGAATAGTAC TACTAACTAA TACTAATATA
4401 TAGAAATAGA TATctagcta gAAATAGAAA CAACTAATAT ATAGATAATC
4451 GAAATTGAAA AGAACTGTCT TTTCTGTATA CTTTGCCCGT TCTATTGCTA
4501 CCGCGGGTCT TATGCAATCG ATCGGATCAT ATAGATATCC CTTCAACACA
4551 ACATAGGTCA TCGAAAGGAT CTCGGACGAC TCACCAAAGC ACGAAAGCCA
4601 GTTAGAAAAT GGATTCCTAT TTGAAGAGTG CCTAACCGCA TGGATAAGCT
4651 CACATTAACC CGTCAATTTT GGATCCAATT CGGGATTTTT CTTGGGAAGT
4701 TTCGGGAAGA AATTGGAATG GAATAATATA GATTCATACA GAGGAAAAGG
4751 TTCTCTATTG ATGCAAACGC TGTACCTAGA GGATAGGGAT AGAGGAAGAG
4801 GGAAAAATCG AAATGAAATA AATAAAGAAT AAAGCAAAAA AAAAATAAGT
4851 CGAAGATAGA AGAGCCCAGA TTCCAAATGA AGAAATGGAA ACTCGAAAAG
4901 GATCCTTCTG ATTCTCAAAG AATGAGGGGC AAGGGGATTG ATACCGAGAA
4951 AGATTTCTTC TTATTATAAG ACGTGATTTG ATCCGCATAT GTTTGGTAAA
5001 AGAACAATCT TCTCCTTTAA TCATAAATGG AAAGTGTTCA ATTAGAACAT
5051 GAAAACGTGA CTCAATTGGT CTTAGTTAGT CTTCGGGACG GAGTGGAAGA
5101 AAGGGCGAAG ACTCTCGAAC GAGGAAAAGG ATCCCTTCGA AAGAATTGAA
5151 CGAGGAGCCG TATTAGGTGA AAATCTCATG TACGATTCTG TAGAGGGACA
5201 GGAAGGGTGA CTTATCTGTC GACTTTTCCA CTATCAACCC CAAAAAACCC
5251 AACTCTGCCT TAC
```

aadA

L79106B

Rice Right targeting sequence

125878

Gene	Product	Plasmid
aadA16gfp	FLARE16-S	pMSK51 (BS)
aadA16gfp-S1	FLARE16-S1	pMSK56 (Nt-pRV111B)
aadA16gfp-S2	FLARE16-S2	pMSK57 (Nt-pRV111B)
aadA11gfp-S3	FLARE11-S3	pMSK49 (Os-pMSK49)

Figure 35

PCT

GENERAL POWER OF ATTORNEY

(for several international applications filed under the Patent Cooperation Treaty)

(PCT Rule 90.5)

The undersigned person(s):
(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)

ADAMS, William T.
Director, Office of Corporate Liaison and Technology Transfer
RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
Old Queens, Somerset Street
New Brunswick, New Jersey 08903
United States of America

hereby appoint(s) the following person as: ☒ agent ☐ common representative

Name and address
(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)

RIGAUT, Kathleen D.
HAGAN, Patrick J.
DORFMAN, John C.
HERRELL, Roger W.
PIPER, Donald R., Jr.
PACE, Vincent T.
SKILLMAN, Henry H.

DANN, DORFMAN, HERRELL AND SKILLMAN
1601 Market Street
Suite 720
Philadelphia, Pennsylvania 19103-2307
United States of America

to represent the undersigned before

☒ all the competent International Authorities

☐ the International Searching Authority only

☐ the International Preliminary Examination Authority only

in connection with any and all international applications filed by the undersigned with the following Office

US/RO as receiving Office and to make or receive payments on behalf of the undersigned.

Signature(s) *(where there are several persons, each of them must sign, next to each signature, indicate the name of the person signing and the capacity in which the person signs, if such capacity is not obvious from reading this power):*

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY

William T Adams

William T. Adams
Director, Office of Corporate Liaison and Technology Transfer

Date: 7-8-98

UTILITY
Original U.S. or PCT D/O

APR 23 2001

DECLARATION, POWER OF ATTORNEY AND POWER TO INSPECT

As a below named inventor, I hereby declare:

that my residence, post office address and citizenship are as stated below next to my name;

that I verily believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural inventors are named below) of the invention entitled: **TRANSLATION CONTROL ELEMENTS FOR HIGH-LEVEL PROTEIN EXPRESSION IN THE PLASTIDS OF HIGHER PLANTS AND METHODS OF USE THEREOF**

the specification of which [check one(s) applicable]

_X_ was filed August 3, 1999 as International Application No. PCT/US99/17806, on which U.S. Patent Application No. 09/762,105 is based.

____ and was amended by Amendment filed __________ (if applicable); [or];

____ is attached to this Declaration, Power of Attorney and Power to Inspect;

that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above; and that I acknowledge my duty to disclose information which is material to the examination of this application in accordance wit Rule 56 (a) [37 C.F.R. §1.56(a)].

CLAIM UNDER 35 USC §119(e): *I hereby claim the benefit under 35 USC §119(e) of any United States provisional applications listed below:*

Provisional Application No.	Filing Date Day/Mo/Year
60/095,163	3 August 1998
60/095,167	3 August 1998
60/112,257	15 December 1998
60/131,611	29 April 1999
60/138,764	11 June 1999

POWER OF ATTORNEY: As inventor, I hereby appoint **DANN, DORFMAN, HERRELL AND SKILLMAN, P.C.** of Philadelphia, PA, and the following individual(s) as my attorneys or agents with full power of substitution to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: **Kathleen D. Rigaut, Ph.D., J.D. Reg. No. 43,047; Maria Kourtakis, Esq. Reg. No. 41,126 and Patrick J. Hagan, Esq. Req. No. 27,643**

3

POWER TO INSPECT: I hereby give **DANN, DORFMAN, HERRELL AND SKILLMAN, P.C.** of Philadelphia, PA or its duly accredited representatives power to inspect and obtain copies of the papers on file relating to this application.

SEND CORRESPONDENCE TO: CUSTOMER NUMBER 000110.

DIRECT INQUIRIES TO: Telephone: (215) 563-4100
Facsimile: (215) 563-4044

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

SOLE OR FIRST JOINT INVENTOR

HW

Full Name ~~Pal~~ Maliga
First Middle Last

Signature [signature]

Date 03/09/2001

Residence East Brunswick New Jersey NJ
City State or Country

Citizenship Hungary

Post Office Address:

72 Yorktown Road

East Brunswick New Jersey 08816
City State or Country Zip Code

SECOND JOINT INVENTOR (IF ANY)

HK

Full Name Hiroshi Kuroda
First Middle ~~Last~~

Signature Hiroshi Kuroda

Date 03/09/2001

Residence Highland Park New Jersey NJ
City State or Country

Citizenship Japan

Post Office Address:

2 Redcliffe Avenue, Apartment 2C

Highland Park New Jersey 08904
City State or Country Zip Code

THIRD JOINT INVENTOR

3(b) Full Name Muhammad Sarwar Khan
First Middle Last

Signature [signature]

Date 04/19/2001

Residence Orlando Florida FL
City State or Country

Citizenship Pakistan

Post Office Address: 2550 N. Alafaya Trail
Apt # 6100 Cottam Landing
Orlando Florida FL-32826
City State or Country Zip Code

FOURTH JOINT INVENTOR

Full Name
First Middle Last

Signature

Date

Residence
City State or Country

Citizenship

Post Office Address:

City State or Country Zip Code

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☐ Page(s)___________of______________________________ were not present for scanning.
(Document title)

☐ Page(s)___________of______________________________ were not present for scanning.
(Document title)

☑ *Scanned copy is best available.* Drawing.